Department of Economics
Issn 1441-5429
Discussion paper 02/11

Mathematical Economics: A Reader

Birendra K. Rai[1], Chiu Ki So[2] and Aaron Nicholas[3]

Abstract:
This paper is modeled as a hypothetical first lecture in a graduate Microeconomics or Mathematical Economics Course. We start with a detailed scrutiny of the notion of a utility function to motivate and describe the common patterns across Mathematical concepts and results that are used by economists. In the process we arrive at a classification of mathematical terms which is used to state mathematical results in economics. The usefulness of the classification scheme is illustrated with the help of a discussion of fixed-point theorems and Arrow's impossibility theorem. Several appendices provide a step-wise description of some mathematical concepts often used by economists and a few useful results in microeconomics.

Keywords. Mathematics, Set theory, Utility function, Arrow's impossibility theorem

[1] Corresponding Author: Department of Economics, Monash University, Clayton, VIC, Australia, Email: birendra.rai@monash.edu.au
[2] Department of Economics, Monash University, Clayton, VIC, Australia.
[3] Graduate School of Business, Deakin University, Burwood, VIC, Australia.

1. Where do economic agents operate?

Let us consider a standard question in consumer theory – What will be the optimal consumption bundle of an agent given her utility function, the amount of money she plans to spend, and the prices of the goods? Formulating this question is a two step procedure. In the first step we translate the intuitive understanding of the consumer's problem into a mathematical framework. In the second step we implicitly argue that the consumer will choose that feasible bundle which offers her maximum utility.

One may say that this question is stated as if the playground is given and we want to predict how the player will play. The crucial thing to note is that mathematics is used, first and foremost, to delineate the precise structure of the playground in which we allow economic agents to operate. Answering *how* an agent will operate is logically the second step. Economists would summarize the answer to this second question by saying that the agent will optimize given the relevant constraints.

We shall be interested in the first step. One of our main goals will be to highlight the common procedure that is used to formulate the playgrounds in which we let economic agents operate. This will allow us to provide a succinct and informative answer to the first question: *Where* do economic agents operate? In order to answer this question we first need to understand the meaning of a utility function.

A. Utility representation theorem

Let us look at what it takes to be able to write down a utility function for a decision maker faced with a finite set of alternatives. We will list each step involved in arriving at the utility representation theorem for this case, describe the meaning of the step, and raise questions about the assumptions in each step. The aim will be to show that a detailed scrutiny of this result (which is one of the first results encountered by students) can be used to motivate and explain the intuition behind mathematical concepts and results.

1. The Unstructured Set: $\mathbb{X}$ is a non-empty finite set of alternatives faced by an agent.[4]
Description: We assume the existence of a set $\mathbb{X}$ that contains a finite number of alternatives from which the agent will have to make a choice. At this stage, the only thing we know about the set X is the elements it contains and the rule for membership in this set. In this case, the rule for membership is that all these elements are being considered by the agent to come up with a final choice. It is also important to note that each element in $\mathbb{X}$ is irreducible in the sense that it can not be broken down any further into sub-elements. We shall refer to a set of irreducible elements as an *unstructured set* if we know nothing more than the identity of the elements it contains. Now let us question the obvious and the not so obvious assumptions we have made in this step.

- What if the set of alternatives is not finite? When do we label a set as a finite set? Do all non-finite sets contain the same number of elements?

[4] In this paper we will neither discuss the pre-requisites of logic (sentential logic, followed by quantifier theory) that are necessary to formulate a theory of sets, nor an axiomatization of set theory.

- What is the nature of the alternatives contained in the set? What if they are consumption streams spread over finite or infinite time horizon, lotteries (von-Neumann and Morgenstern, 1944), or acts whose outcomes depend on the state of world which is uncertain at the moment of making the decision (Savage, 1954)? Will these different possibilities require completely different approaches to come up with a utility representation theorem?

- Does every collection of objects qualify as a set?[5]

2. The Tool: R is a binary relator (or, a binary relation) defined over the elements of $\mathbb{X}$ where xRy means the agent believes 'x is at least as good as y.'
Description: We assume that the agent possesses a tool - a binary relator - that will (potentially) allow her to establish a relationship between any pair of elements in the set $\mathbb{X}$ (and thereby provide some structure to the elements of $\mathbb{X}$). Given the nature of the issue under consideration, we endow the tool with a specific meaning to make precise the nature of the relationship the agent is assumed to establish between any pair of elements from the set $\mathbb{X}$. In this context, we assume the tool stands for 'is at least as good as.' The following questions immediately come to mind.

- Why should we assume that the agent uses one, and only one, binary relator to structure the unstructured set of alternatives. For instance, it seems no less reasonable to assume that one might use two binary relators in a sequential manner (Manzini and Marriotti, 2007; Rubinstein and Salant, 2008). The first binary relator could help her *partition* the set of alternatives into two disjoint subsets: the first subset containing all those elements the agent thinks she may ultimately choose, and the second containing those elements the agent is sure she will definitely not choose. The second binary relator may then be used to *rank* only the elements of the first subset.

- Is the binary relator a 'new' entity or is it derived from the unstructured set of alternatives we started with?

3.1 Properties of the Tool: R is a weak-order over $\mathbb{X}$. In other words, the binary relator satisfies the properties of transitivity and completeness.
Description: We assume this tool satisfies two properties that we consider to be reasonable in this context. Transitivity requires that if the agent believes x is at least as good as y and y is at least as good as z, then she must also believe that x is at least as good as z. Completeness imposes the restriction that the agent must be able to compare every pair of alternatives in the unstructured set $\mathbb{X}$.

- There is nothing that forces the decision maker to have a transitive and complete binary relator. These are assumptions we have made. Hence, we can question these properties and replace them with other properties if intuition and/or empirical evidence suggests so (Gilboa, 2009). What could be those other properties?

[5]We can do no better than point the reader to Nagel and Newman (2008) for a discussion of the importance of this question in modern mathematics.

3.2 The Structured Set: $(\mathbb{X}, R)$ is a weakly-ordered set.

Description: Transitivity and completeness of the binary relator ensure that the agent can rank all the alternatives from the most preferred to the least preferred (she may give the same rank to two distinct alternatives). We are referring to the set of alternatives now as a *structured set* because this is precisely what the tool allows the agent to do: provide some structure to the initially unstructured set of alternatives. Given the nature of the tool and its properties, in addition to knowing what are the irreducible elements of $\mathbb{X}$, we also have the information that the agent will have a ranking over the elements of the set. Given an unstructured set, if we use a binary relator as the tool and endow it with the properties of transitivity and completeness, then the structured set we obtain is referred to as a *weakly-ordered set*.

- Will we get a different structured set if we replace transitivity or/and completeness with some other properties?
- We started with the unstructured set of alternatives and assumed the decision maker uses a binary relator satisfying two properties to convert the unstructured set of alternatives into a structured set. Do we obtain all structured sets via the same procedure?
- Are binary relators the only tools that can be used to structure unstructured sets? If there are other tools, are they fundamentally different from binary relators? Or, are all tools expressions of a single underlying concept?

4. Utility Representation Theorem: It may be stated as follows.

If

- $\mathbb{X}$ is a finite unstructured set of alternatives; and,
- the binary relator R defined over $\mathbb{X}$ is transitive and complete,

then

- there exists a real-valued function $u : \mathbb{X} \to \mathbb{R}$ such that xRy if and only if $u(x) \geq u(y)$.

Description: If the binary relator satisfies properties that help structure the unstructured set of alternatives into a weakly ordered set, then we can write down a utility function for the agent. In other words, we can *map* each alternative in the set of alternatives to an element in the set of *real number* such that if the agent thinks alternative x is at least as good as alternative y, then the real number associated with x will be at least as large as the real number associated with y. The list of questions one could ask at this step is quite long. In this paper, we wish to raise only one.

- What are real numbers?

Figure 1: The Common Procedure

2. Converting an unstructured set into a structured set

The set of real numbers is a *structured set*. So is a pre-ordered set, a partially ordered set, a weakly ordered set, a linearly ordered set, a lattice, a group, an ordered field, a metric space, a vector space, a normed vector space, an inner product space, a Banach space, a Polish space, a Hilbert space, a topological space, a probability space, and ... They are often used in economic modeling and most of them appear in the first year graduate textbooks by Mas-Colell et al. (1995), Myerson (1997), and Stokey and Lucas (1989).

Structured sets are usually the playgrounds where economic agents operate in mathematical models. Instead of highlighting the differences between these structured sets, we want to focus on the one feature they all have in common. They are all formulated by using, essentially, a *common procedure*: we start with an unstructured set, choose some tools, and endow these tools with certain properties to add some structure to the elements of the unstructured set (Figure 1).

Starting with the same unstructured set we end up with different structured sets if the tools and/or their properties differ. We next provide a brief outline of this common procedure. Even the readers familiar with Axiomatic Set Theory (Bourbaki, 1950; Debreu, 1959; Halmos, 1974; Weintraub, 2002) may find the exposition interesting.

The notion of an unstructured set is the starting point of our discussion (please refer to Table 1 for the technical notation used throughout the paper). We can talk about some notions related to sets in a meaningful way even if we know nothing more than the irreducible elements of the set(s) we are considering. Subset of a set, cardinality of a set, power set of a set, union and intersection of two sets, difference of two sets, disjoint subsets of a set, partition of a set, cartesian product of two sets, and n-fold cartesian product of a set are examples of such notions.[6]

Before proceeding we would like to elaborate on the notions of 'power set' of a set and the 'cartesian product set' of (two or more) sets as they will be appear quite often in the following discussion. Consider set $\mathbb{X} = \{x_1, x_2\}$ and set $\mathbb{Y} = \{y_1, y_2, y_3\}$. The power set of

[6] It is crucial to note that it takes certain axioms of set theory to assert the existence of, say, a power set of any given set or the union of two sets. Moreover, one may use different combinations of axioms to formulate a theory of sets. Zermelo-Fraenkel set theory with the axiom of choice (ZFC) is the dominant framework used by economists. Several researchers have questioned the reliance on ZFC to mathematically represent economic notions because of its non-constructivist features: the existence of some mathematical objects is assumed in ZFC even if one can not provide an algorithm to show how those objects may be constructed. A detailed discussion of these issues is beyond the scope of the present paper. The interested reader may refer to Vellupillai (2010) for a critical discussion of the drawbacks of ZFC and suggestions for a constructive framework for economic modeling.

Table 1: Useful Notation†

Symbol	**Meaning**
$\mathbb{X} = \{x_1, x_2\}$	an unstructured set whose irreducible elements are x_1 and x_2
$\emptyset$	the empty set, or null set (a set that contains no elements)
$\lvert\mathbb{X}\rvert$	Cardinality of a set (number of elements in the set)
$(\mathbb{X}, T)$	Structured set (where T is the tool used to add structure to $\mathbb{X}$)
$\in$	belongs to (for example, $x_1 \in \mathbb{X} = \{x_1, x_2\}$)
$\forall$	for all (for example, $x \geq 3,\ \forall x \in \mathbb{X} = \{3, 4, 5\}$)
$\exists$	there exists (for example, $\exists\ x \in \mathbb{X} = \{3, 4, 5\}$ such that $3 < x < 5$)
$\subset$	is a proper subset of (for example, $\{x, y\} \subset \mathbb{X} = \{x, y, z\}$)
$P(\mathbb{X})$	Power-set of set $\mathbb{X}$ (the collection of all possible subsets of $\mathbb{X}$)
$\mathbb{X} \times \mathbb{Y}$	Cartesian product of set $\mathbb{X}$ and set $\mathbb{Y}$
$\mathbb{Y} \times \mathbb{X}$	Cartesian product of set $\mathbb{Y}$ and set $\mathbb{X}$
$\mathbb{X} \times \mathbb{X}$	2-fold Cartesian product of set $\mathbb{X}$ (also denoted as $\mathbb{X}^2$)

† See the Appendix for a detailed example.

a set contains all possible subsets of the given set (including the empty set). Thus,

$$P(\mathbb{X}) = \{\emptyset, \{x_1\}, \{x_2\}, \{x_1, x_2\}\}, \text{ and}$$

$$P(\mathbb{Y}) = \{\emptyset, \{y_1\}, \{y_2\}, \{y_3\}, \{y_1, y_2\}, \{y_1, y_3\}, \{y_2, y_3\}, \{y_1, y_2, y_3\}\}.$$

The cartesian product of set $\mathbb{X}$ and set $\mathbb{Y}$ is given by

$$\mathbb{X} \times \mathbb{Y} = \{(x_1, y_1), (x_1, y_2), (x_1, y_3), (x_2, y_1), (x_2, y_2), (x_2, y_3)\}.$$

We should think of $\mathbb{X} \times \mathbb{Y}$ as *one* set. The six elements listed in the set above are the irreducible elements of 'the' set $\mathbb{X} \times \mathbb{Y}$. Each irreducible element of the set $\mathbb{X} \times \mathbb{Y}$ is an *ordered* pair composed of two entities listed in the order in which the sets are multiplied: the first is an irreducible element of the set $\mathbb{X}$ and the second is an irreducible element of the set $\mathbb{Y}$. The cartesian product set of n sets will be a collection of *ordered n-tuples*.

We can similarly formulate the n-fold cartesian product set of a given set. For example, the 2-fold cartesian product set of set $\mathbb{X}$ is given by

$$\mathbb{X} \times \mathbb{X} = \{(x_1, x_1), (x_1, x_2), (x_2, x_1), (x_2, x_2)\}.$$

The set $\mathbb{X} \times \mathbb{X}$ may also be denoted as $\mathbb{X}^2$. The irreducible element (x_1, x_2) is considered distinct from the irreducible element (x_2, x_1). We may imagine that there are two copies of the set $\mathbb{X}$ and we keep track of whether an element came from the first copy or the second copy of set $\mathbb{X}$.

Suppose we start with *one* unstructured set. All the tools that can help us convert this unstructured set into a structured set are expressions of a single concept (which will be discussed later). In practice, it is convenient to attach a label to a tool in order to clarify what the tool allows us to do with/to the elements of the unstructured set.

The two basic tools used to structure unstructured sets are labeled as *binary relators* and *binary operators*. Binary relators help us talk about a certain type of relationship between two elements of a given set. For example, the binary relator 'is at least as good as' is often used to talk about the preferences of an agent over a set of alternatives. Given a set a people, 'is in love with', 'is married to', 'is taller than' are some examples of binary relators that can be used to describe a particular type of relationship between two people belonging to the set. In contrast, a binary operator uses a pair of elements belonging to a set as input and produces a third entity belonging to the same set as the output.[7]

A. Binary relators as structuring tools

Given an unstructured set $\mathbb{X}$, a *binary relator R defined over set* $\mathbb{X}$ is a tool that can be used to describe the relationship between two irreducible elements of $\mathbb{X}$. If x and y are irreducible elements of the unstructured set $\mathbb{X}$, then xRy is read as 'x is related to y.' It

[7] At this stage, one may think of addition and multiplication over the set of real numbers as examples of binary operators.

is crucial to note that the notion of a binary relator can be invoked only after we have specified the unstructured set. Let us now look at some properties that binary relators can be endowed with (see Tables 2 and 3).

Reflexivity: A binary relator R defined over a set $\mathbb{X}$ is reflexive if xRx is true for every x in the set $\mathbb{X}$. We will accept the statement 'person x is as tall as person x.' But, will we accept the statement – x is a son of x – for any human being? These examples highlight that reflexivity is not a vacuous property (an impression one may get in the context of individual choice theory where it seems superfluous to say that object x is at least as good as object x from the perspective of an individual).

Transitivity: A binary relator R defined over a set $\mathbb{X}$ is transitive if xRy and yRz imply that xRz is true. If x is an enemy of y, and y is an enemy of z, then our experience with realpolitik might suggest that x may not be an enemy of z. Thus, when the unstructured set is a group of people, the binary relator 'is an enemy of' may not be transitive. On the other hand, the binary relator 'is at least as heavy as' will be transitive.

Completeness: A binary relator R defined over a set $\mathbb{X}$ is complete if xRy or/and yRx for any x, y in the set $\mathbb{X}$. It requires that we should be able to relate any pair of elements from the unstructured set. Once again, consider a decision maker faced with an unstructured set of alternatives $\mathbb{X} = \{a, b, c\}$, where the alternatives are three job offers she has. Assume that she uses the binary relator 'at least as good as' to relate the elements of the unstructured set. Even after days of thinking the decision maker may not be able to conclude either bRc or cRb. This does not mean that she is indifferent between the two jobs. It could really mean that she can not compare the two. When we assume completeness, we rule out such a possibility. If we set $y = x$, then completeness will tell us that xRx. Hence, if a binary relator defined over some set is complete, then it will also be reflexive.
Symmetry: A binary relator R defined over a set $\mathbb{X}$ is symmetric if xRy implies that yRx is true. Consider the binary relator R with the meaning 'is married to.' For a given set of people, if we know x is married to y, then our experience with our world suggests that y is married to x. Thus, 'is married to' is a symmetric binary relator over a set of people. Now consider the binary relator 'is in love with' over the same set of people. Clearly, even if it is true that x is in love with y, it does not necessarily imply that y is in love with x. Hence, the binary relator 'is in love with' may or may not be symmetric over a given set of people.

Asymmetry: A binary relator R defined over a set $\mathbb{X}$ is asymmetric if xRy implies that yRx is false. Consider the binary relator R with the meaning 'is a son of.' If we know that x is a son of y, then we can conclude that y can not be a son of x. Thus, 'is a son of' is an asymmetric binary relator over a set of people.

Antisymmetry: Consider the set of real numbers and assume the binary relator R stands for 'is weakly greater than.' $xR3$ would mean that the real number x is weakly greater than 3. $3Rx$ would mean that 3 is weakly greater than x. Let us now try to find all the real numbers x such that both $xR3$ and $3Rx$ are true. The only number x that will allow both these conditions to hold is 3. In other words, antisymmetry requires that if x is related to y and y is related to x, then x and y must be the same irreducible element

Table 2: Some Properties of Binary Relators

Property of R	**Meaning**
Reflexivity	For every x in set $\mathbb{X}$, xRx is true.
Transitivity	For every x, y, z in set $\mathbb{X}$, if xRy and yRz are true, then xRz is true.
Completeness	For every x, y in set $\mathbb{X}$, either xRy or yRx (or, both).
Symmetry	For every x, y in set $\mathbb{X}$, if xRy is true, then yRx is true.
Asymmetry	For every x, y in set $\mathbb{X}$, if xRy is true, then yRx is false.
Antisymmetry	For every x, y in set $\mathbb{X}$, if xRy and yRx are true, then $x = y$.

Table 3: Binary relators defined over a set of people

Property of R	**Example**	**(Possible) Counterexample**
Reflexivity	is at least as tall as	is the son of
Transitivity	is at least as heavy as	is an enemy of
Completeness	is at least as heavy as	is strictly heavier than
Symmetry	is married to	is in love with
Asymmetry	is the mother of	is the brother of
Antisymmetry	has the fingerprint of	is as tall as

Table 4: Structured sets using one binary relator

Properties of the Binary relator	Structured Set
Reflexivity, Transitivity	*Pre-ordered set*
Reflexivity, Transitivity, Antisymmetry	*Partially-ordered set*
Transitivity, Completeness	*Weakly-ordered set*
Transitivity, Completeness, Antisymmetry	*Linearly-ordered set*

in the unstructured set. Now consider a decision maker faced with an unstructured set of alternatives $\mathbb{X} = \{a, b, c\}$, where the three elements are an apple, a banana, and a carrot. Assume that she uses the binary relator 'at least as good as' to relate the elements of the unstructured set. Suppose she feels aRb and bRa. Or, she feels that an apple is at least as good as a banana, and a banana is at least as good as an apple. This in no way implies that an apple is the same object as a banana. In short, antisymmetry rules out indifference between distinct elements of the set.

Table 4 provides a list of some structured sets that can be formulated using *one* binary relator as the structuring tool. For instance, if we use a binary relator satisfying reflexivity and transitivity, then the resulting structured set is labeled a pre-ordered set.

When we are doing pure mathematics we do not need to specify the *meaning* of the binary relator. We may simply ask what type of structured set will emerge if we endow the tool with certain properties. In contrast, when we apply mathematics to any 'real world' issue then we have to specify the meaning of the tools we are using. For instance, if a sociologist wants to study the father-son relationship over a given set of men, then she will have to specify the meaning of the binary relator as 'is the son of.' Moreover, once we specify the meaning of the binary relator, we loose the freedom to endow it with any property we want if we want to make logically valid statements. Given a set of men, we will accept the statement 'person x is as tall as person x.' But, we will not accept the statement 'x is the son of x.' The binary relator 'is as tall as' can be endowed with the property of reflexivity but 'is the son of' can not.

We rely on our experiences with our world while making the judgement that 'x is a son of x' is not a valid statement. Thus, pure mathematics allows us to conceive worlds other than the one we inhabit. One such world which is mathematically feasible but different from ours would be where the statement 'x is a son of x' is perhaps true.

B. Binary operators as structuring tools

Given an unstructured set $\mathbb{X}$, a *binary operator defined over set* $\mathbb{X}$ can be thought of as a machine that takes any pair of irreducible elements from $\mathbb{X}$ as the input and produces an irreducible element belonging to $\mathbb{X}$ as the output. Let $x \diamond y$ denote the output when we first feed x and then y into the binary operator denoted by $\diamond$. Once again, the notion of a binary operator can be introduced only after we have specified the set over which it will operate. Table 5 summarizes some properties that binary operators can be endowed with.

Closure: If x and y are any two elements from the set $\mathbb{X}$, then $x \diamond y$ must also belong to set $\mathbb{X}$. In words, closure requires that the output produced by the binary operator must be an irreducible element of the unstructured set we started with. Consider 'adding' two natural numbers contained in the set of natural numbers. The final outcome is also a natural number. But, subtracting two natural numbers may lead to an output that does not belong to the set of natural numbers. Note that we have included the notion of closure in the definition of a binary operator.

Associativity: The binary operator, as per our definition, can act on only two elements at a time. Suppose we have three elements that have to be used to produce a final output. We will necessarily have to first use two out of these three elements to produce an intermediate output; and then combine this intermediate output with the remaining element to obtain the final output. Suppose we lay down the three elements in the sequence $x-y-z$ on a table from left to right. We have to first pick any two adjacent elements. We have only two possibilities: x and y, or y and z. Suppose we pick x and y. Associativity then asks us to feed x and then y into the binary operator and collect the intermediate output $w = x \diamond y$. Then, we must feed w followed by z, respecting the order in which we had laid the three elements on the table. If we had instead picked y and z, then we would have to first feed y and then z into the binary operator. Let the resulting intermediate output be $v = y \diamond z$. In order to get the final output in this case, we will have to feed first x and then v. Associativity requires that the final output in the two cases must be the same.

Remark 1. The more changes we do to any given production process, the more likely it will be that the output will vary. Associativity develops a notion of what would be a minor perturbation of the production process. It requires the same output if the production process is perturbed in this minor way.

Existence of a unique identity element with respect to the binary operator: It stipulates the existence of a special element in the unstructured set. It says that there must exist a special irreducible element in the unstructured set such that whenever we feed this special element and another element (which may be this special element itself), then we obtain the other element as the output irrespective of the order in which we feed the two elements into the binary operator. This should remind us of the role played by *zero* with respect to addition, or of the role played by *one* with respect to multiplication, over the set of real numbers. This in turn clarifies why we have denoted the identity element as $s_\diamond$ and not simply s. The subscript is used to highlight that there could be several binary operators that satisfy associativity, and each such operator may induce a different identity element.

Table 5: Some Properties of Binary Operators

Property of $\diamond$	**Meaning**
Closure	$x \diamond y \in \mathbb{X}, \ \forall x, y \in \mathbb{X}.$
Associativity	$(x \diamond y) \diamond z \ = \ x \diamond (y \diamond z), \forall x, y, z \in \mathbb{X}.$
Identity	$\exists \ s_\diamond \in \mathbb{X}$ such that $s_\diamond \diamond x \ = \ x \diamond s_\diamond \ = \ x, \forall x \in \mathbb{X}.$
Inverse	If $x \in \mathbb{X}$, then $\exists \ \tilde{x}_\diamond \in \mathbb{X}$ such that $x \diamond \tilde{x}_\diamond = \tilde{x}_\diamond \diamond x = s_\diamond.$
Commutativity	$x \diamond y = y \diamond x, \forall x, y \in \mathbb{X}.$

To repeat, addition and multiplication both satisfy associativity over the set of real numbers, but the identity element induced by addition is zero, whereas the identity element induced by multiplication is one.

Existence of an inverse element for every element in the unstructured set with respect to the binary operator: Recall our intuitive understanding of addition over the set of real numbers. For each real number x in the set of real numbers, there exists the real number $-x$ in the set of real numbers, such that if we add the two together the outcome is zero, irrespective of the order in which we add the two numbers. Moreover, the output (zero) is the identity element with respect to addition over the set of real numbers. Similarly, for each real number $x \neq 0$ in the set of real numbers, there exists the real number x^{-1} in the set of real numbers, such that if we multiply the two together the outcome is one, irrespective of the order in which we multiply the two numbers. Moreover, one is the identity element with respect to multiplication over the set of real numbers. The inverse property captures these intuitive notions.

Commutativity: If x and y are any two elements from the set $\mathbb{X}$, then $x \diamond y$ must be the same as $y \diamond x$. Or, it requires that the output produced by the binary operator should not depend on the order in which the two inputs are fed into the operator. Addition over the set of real numbers is commutative but subtraction is not.

Remark 2. Mathematics is often regarded as a language that helps us structure our discourse about real world phenomena. We use the word 'add' in conversations in our daily lives in numerous contexts. In many of these contexts the order in which two things are 'added' to produce a third thing does matter. For instance, a common way to cook pasta is to first add water in a pan and heat it for, say, fifteen minutes, and then add raw pasta into the hot water. The output is likely to be different if we reverse the order by first adding raw pasta to the pan, heating it for fifteen minutes, and then adding water.

Table 6: STRUCTURED SETS USING ONE BINARY OPERATOR

Properties of the Binary Operator	**Structured Set**
Closure	*Magma*
Closure, Associativity	*Semi-group*
Closure, Associativity, Identity	*Monoid*
Closure, Associativity, Identity, Inverse	*Group*
Closure, Associativity, Identity, Inverse, Commutativity	*Abelian Group*

As with binary relators, when we are doing pure mathematics we can reasonably ask what type of structured set will emerge if we endow a binary operator with certain properties. For instance, a *Group* is a structured set obtained by endowing a binary relator with all the properties listed above except commutativity (see table 6). Some of the structured sets listed in Table 6 are rarely explicitly encountered in economic modeling. We have listed them solely to highlight that the same procedure is used to generate all of them.

We may also use two or more tools to provide some structure to the elements of an unstructured set. For instance, we need to use three tools (the binary relator 'is weakly greater than' and the binary operators of addition and multiplication) to arrive at the structured set referred to as the set of rational numbers. The set of real numbers is obtained by imposing an additional property.[8]

C. Tools identify subsets

We have looked at two tools that can be used to add structure to an unstructured set. A natural question to ask would be: Can these seemingly distinct tools be understood as expressions of a *single* concept? These tools *identify* a subset of a suitably defined set derived from the unstructured set we start with. In fact, one may say that tools *are* subsets.

C.1. Binary relator defined over $\mathbb{X}$ identifies a subset of $\mathbb{X}^2$

Consider the set $\mathbb{X} = \{a, b, c\}$, where the three elements refer to an apple, a banana, and a carrot. The 2-fold cartesian product of $\mathbb{X}$ with itself will be

$$\mathbb{X}^2 = \{(a,a),(a,b),(a,c),(b,a),(b,b),(b,c),(c,a),(c,b),(c,c)\}.$$

Suppose a decision maker has a binary relator R defined over the set $\mathbb{X}$.

[8]See the Appendix for details.

Case 1. Let R stand for 'is at least as good as.' Suppose we know that R is reflexive. This would mean that aRa, bRb, and cRc are true. The one and only subset of $\mathbb{X}^2$ identified by this R is

$$\mathbb{A}_1 \;=\; \{(a,a),(b,b),(c,c)\}.$$

Case 2. Let R stand for 'is strictly better than.' Now suppose we know that R is transitive and complete and aRb, bRc, and aRc are true. We may say that this R identifies the set

$$\mathbb{A}_2 \;=\; \{(a,b),(b,c),(a,c)\}.$$

Remark 3. Case 2 highlights the importance of treating the irreducible elements of $\mathbb{X}^2$ as *ordered* pairs. Specifically, the R in Case 2 identifies one and only one subset of $\mathbb{X}^2$ (i.e., $\mathbb{A}_2$). It does not identify, for instance,

$$\mathbb{A}_3 \;=\; \{(b,a),(b,c),(a,c)\} \;\text{ or }\; \mathbb{A}_4 \;=\; \{(b,a),(c,b),(c,a)\}.$$

C.1. Binary operator defined over $\mathbb{X}$ identifies a subset of $\mathbb{X}^3$

Consider an unstructured set $\mathbb{X}$ containing countable elements $\{\ldots,\alpha,\beta,\gamma,\ldots\}$. First note that the irreducible elements of the 3-fold cartesian product of $\mathbb{X}$ with itself – $\mathbb{X}^3$ – will be ordered triples of the form (β,γ,α), (α,γ,β), etc. We can interpret the ordered triple (α,β,γ) as saying: when we first feed α and then β into the binary operator, it produces γ as the output.

Now consider the process of using a binary operator (denoted by the symbol $\diamond$) to provide some structure to $\mathbb{X}$. We have assumed that binary operators satisfy the closure property, by definition. Consequently, if we feed any two irreducible elements of $\mathbb{X}$ into the binary operator $\diamond$, then the output will also be an irreducible element of $\mathbb{X}$. The binary operator $\diamond$ identifies a subset $\mathbb{S}$ of $\mathbb{X}^3$.

To understand this intuitively, let us look at the set of all integers $\mathbb{Z}$ and treat the binary operator $\diamond$ as the standard operator for addition. We know that $(-3)+5=2$. This fact can be represented as the ordered triple $(-3,5,2)$ if we agree that the rule we will use to interpret this triple will be as follows: $(-3,5,2)$ means that when we first feed -3 and then 5 into the binary operator $+$, then it produces 2 as the output.

Given this rule of interpretation, the ordered triples $(2,9,-3)$ and $(100,0,0)$ do not make sense. But, ordered triples like $(100,0,100)$ and $(0,0,0)$ do make sense. Clearly, some irreducible elements of $\mathbb{Z}^3$ are consistent with the binary operation of addition, while others are not. This clarifies why we say that a binary operator defined over a set identifies 'a subset of the 3-fold cartesian product of the set with itself.'

This discussion leads us to take note of an important point. The starting point of our discussion is the notion of an unstructured set, say $\mathbb{X}$. We talked about tools that can be used to add structure to $\mathbb{X}$. We may attach a label to the tools – binary relators, binary operators, etc. – but the tools will turn out to be subsets of suitably defined sets derived from $\mathbb{X}$.

Remark 3. Any mathematical notion that we could potentially discuss after specifying the unstructured sets we start with will *necessarily* be a subset of some cartesian product set that can be formulated using the sets we start with. This should not be surprising since the universe of our discourse is limited by the unstructured sets we start with.

3. Dealing with two sets: Mappings

Till now we have been concerned with just one set and the discussion has focused on describing the common procedure that is used to convert an unstructured set into a structured set. The concept of a mapping (a function or a correspondence) comes into the picture when we want to mathematically express a piece of information that involves reference to two sets (where each of them may be unstructured or structured and the two sets may or may not be identical to each other).

Consider two unstructured sets $\mathbb{X}$ and $\mathbb{Y}$ containing the names of all married men and women in a society, respectively. Suppose the society is *monogamous* and only allows heterosexual marriages. Then we can define a function – $f : \mathbb{X} \to \mathbb{Y}$ – which takes the name of a man as the input and produces the name of his wife as the output. Let $f(x)$ denote the element of set $\mathbb{Y}$ that is produced as the output when the element x belonging to set $\mathbb{X}$ is used as the input. We can describe the relationship between the elements of these two sets with the help of a binary relator R with the meaning 'is the husband of.' Clearly, $xRf(x)$ can be read as 'man x is the husband of *woman* $f(x)$.'

For a *polygynous* society we can define a correspondence – $f^c : \mathbb{X} \to \mathcal{P}(\mathbb{Y})$ – which takes the name of a man as the input and produces the names of all his wives as the output. The set $\mathcal{P}(\mathbb{Y})$ contains all the non-empty subsets of the set $\mathbb{Y}$. Thus, a correspondence is a mapping that describes the relationship between an element of one set and at least one element of another set. We may now read $xRf^c(x)$ as 'man x is the husband of *women* $f^c(x)$.'

A. Mappings as structuring tools

Suppose $\mathbb{X} = \{c_1, c_2, c_3\}$ is the set containing the names of the three most populated cities in the world at a given time. We are interested in defining a function that takes any city from the set $\mathbb{X}$ as the input and produces a real number denoting the population of the city as the output. Formally, this function may be denoted as $f : \mathbb{X} \to \mathbb{R}$.

Now suppose we are interested in defining a function that takes any pair of cities from the set $\mathbb{X}$ as the input and produces a real number denoting the distance between the two cities as the output. Formally, this distance function may be denoted as $d : \mathbb{X}^2 \to \mathbb{R}$. The notation used for the input set – $\mathbb{X}^2$ – clarifies that the distance function takes *one* irreducible element from the 2-fold cartesian product of set $\mathbb{X}$ as the input. Note that the set $\mathbb{X}^2$ is a collection of nine irreducible elements. Or,

$$\mathbb{X}^2 = \{(c_1, c_1), (c_1, c_2), (c_1, c_3), (c_2, c_1), (c_2, c_2), (c_2, c_3), (c_3, c_1), (c_3, c_2), (c_3, c_3)\}.$$

If we use any of the nine irreducible elements of $\mathbb{X}^2$ as the input, the function f produces a real number as the output. The world around us does not come equipped with a notion of distance between two points. We choose to define the distance between two points and thus we may do so in many ways. But, we should first ask: What should be the essential properties of a function $d : \mathbb{X}^2 \to \mathbb{R}$ that calculates the distance between two points? We may use the following three properties.

- *Non-negativity*: $d(x, y) \ \geq \ 0$, for all x, y in $\mathbb{X}$, and $f(x, y) \ = \ 0$ iff $x = y$.
- *Symmetry*: $d(x, y) \ = \ d(y, x)$, for all x, y in $\mathbb{X}$.
- *Triangular inequality*: $d(x, y) \ \leq \ d(x, z) + d(z, y)$, for all x, y, z in $\mathbb{X}$.

Once we endow d with these three properties, we obtain one of the most widely used structured sets in economics. The function d is termed a metric and the structured set $(\mathbb{X}, d)$ is called a *Metric Space*. There can be many functions that satisfy the three properties mentioned above. For instance, suppose $\mathbb{X}$ is $\mathbb{R}^2$, i.e., the 2-fold cartesian product of the set of real numbers. Thus, any irreducible element of $\mathbb{X}$ will be a pair of two real numbers. We could calculate the distance between $x, y \in \mathbb{R}^2$ by using

$$d_e(x, y) \ = \ \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2},$$

or some other function that satisfies the three properties. Given any metric space, once we specify a particular way of calculating the distance, we add more structure to it. When we use the metric d_e over $\mathbb{R}^2$, we get the 2-dimensional *Euclidean Space*. This discussion highlights that a function may also be used as a tool to add more structure to a set. As mentioned above (Remark 3), like binary relators and binary operators, functions also identify subsets of suitably defined sets formulated from the sets we start with.

B. Mappings identify subsets

Any function $f : \mathbb{X} \to \mathbb{Y}$ identifies a subset of the set $\mathbb{Y} \times \mathbb{X}$. Consider the exercise of *drawing* the function $f(x) = x^2$ (where f takes a real number as the input and produces a real number as output) on a two dimensional graph with values of x on the horizontal axis and the values of $f(x)$ on the vertical axis. In the first step we draw the horizontal and the vertical axes. It is crucial to recall that this simple act of drawing the two axes amounts to laying out all the elements of the set $\mathbb{R} \times \mathbb{R}$, where $\mathbb{R}$ is the set of real numbers. Every point on this two-dimensional plane is an ordered pair labeled as (y, x), and thus an irreducible element of the set $\mathbb{R} \times \mathbb{R}$. In the next step, we draw the actual function $f(x) = x^2$. The act of drawing the function is nothing more than pointing out the subset of the set $\mathbb{R} \times \mathbb{R}$ that this function identifies.

4. A possible classification of mathematical terms

Sets, binary relators, binary operators, and functions are often regarded as the four basic concepts of mathematics (Gowers, 2008). The discussion of the utility representation theorem presented earlier not only involves a reference to all these concepts but it also suggests a way to classify mathematical terms. This exercise may be a useful way to highlight the intuitive connections between seemingly distinct mathematical concepts.

As mentioned earlier, if Axiomatic Set Theory is used as the framework to formalize economic notions then the starting point of the discourse will necessarily be one or more unstructured sets. All subsequent mathematical terms can then be expressed as subsets of appropriately defined sets derived from the unstructured sets we start with. This is an important observation but may not be of much practical help for an economist. We use the utility representation theorem as a guide to classify mathematical terms into four classes of basic mathematical objects (and concepts/properties related to these basic objects):

- unstructured sets
- structuring tools
- structured sets
- mappings between two sets.[9]

This classification stresses that mathematical modeling in economics often involves using one or more unstructured sets as the building blocks for everything that follows. Figure 2 illustrates this classification scheme. The solid (broken) arrows represent a mapping between one unstructured set and one structured set (two unstructured or two structured sets). The direction of the arrows is being used to distinguish the input set and output set associated with the mapping.

We can talk about the notion of a subset of a set whether it is unstructured or structured. But, we can not talk about the convexity of an unstructured set. An unstructured set, as the name suggests, does not have the structure that will allow us to define the notion of a 'line segment' which is required to talk about the convexity of a set.

On the other hand, a *vector space* is one of many examples of structured sets that have enough structure to allow for a meaningful definition of a line segment and thus we can talk about the notion of a convex subset of a vector space (see the appendix). In general, any concept that can be meaningfully discussed in the context of a given set when is has relatively less structure can also be discussed when it has relatively more structure.

We have already discussed some important properties of the two important tools (binary relators and operators). Functions of various types play an important role in economic analysis. Once again, there needs to be a certain minimal amount of structure in the input and output sets associated with the function so that we may talk about the certain properties such as continuity, convexity, and quasi-concavity.

It is also important to take note of the basic object whose property is being discussed in any mathematical statement. For instance, the conceptual content of 'completeness of a binary relator' differs from that of 'completeness of a metric space.' Table 7 provides a brief summary of this discussion.

5. Why this classification?

Multiple classification schemes can be used to organize mathematical terms. We have tried to strike a balance so that the scheme is neither too abstract to be of little practical use for a student of economics, nor too detailed to distract from the intuitive connections

[9] Although mappings can be expressed as binary relators, we follow the standard convention and treat them as distinct objects.

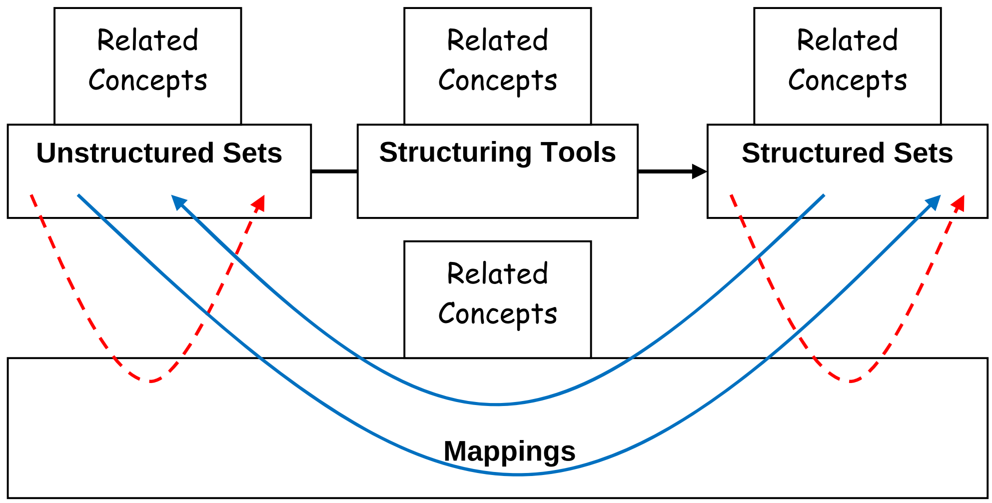

Figure 2: A classification of mathematical terms.

Table 7: A classification of mathematical terms

Basic Object	**Example**	**Some related concepts/properties**
[1] *Unstructured Sets*	–	Cardinality, Subset, Power set
[2] *Structuring Tools*	Binary relator	Transitivity, Symmetry, Completeness
[3] *Structured Sets*	Metric space	Compactness, Connectedness
[4] *Mappings*	Correspondence	Continuity, Convexity, Convex-valued

between various mathematical terms. In this section we shall first illustrate that mathematical results are usually stated using the classification scheme presented above. Thus, familiarity with this classification scheme may be of immense help for students in understanding mathematical results. Moreover, keeping track of the four basic mathematical objects (and their properties) provides a useful framework to organize the *theoretical* literature on economic issues. This will be illustrated with the help of a detailed discussion of Arrow's impossibility theorem.

A. Patterns in mathematical results

The utility representation theorem we have discussed is silent about the nature of alternatives faced by the decision maker. They could be perishable objects for immediate consumption, consumption streams spread over finite or infinite time horizon, lotteries, or acts whose outcomes depend on the eventual state of world which is uncertain at the time of making the decision. There are utility representation theorems for all these, and several other, types of alternatives. If all these theorems are referred to as utility representation theorems, then one would expect some common pattern across them. Similarly, we would expect some common pattern across all mathematical results referred to as fixed point theorems, all results referred to as separating hyperplane theorems, or any other class of results. If we take this chain of thought to its extreme, then we are led to ask whether there is a common pattern across mathematical results, in general.

Mathematical results are usually expressed in the form 'if A, then B.' This statement is too abstract to help us in any significant way. It is almost tautological to say that if there is a pattern across mathematical results, then it must be in the contents of A and B. Having a classification of mathematical terms in mind is useful because it determines the language we use to express the possible forms that the contents of A and B can take in any result. We next use fixed point theorems (which have been used to prove some of the most fundamental results in economic theory) to illustrate that the language of mathematical results employs the above mentioned classification scheme.

Consider the function $f(x) = (x - \frac{1}{2})^2$, where $x \in [0, 1]$. It may be easily verified that the values of $f(x)$ will also lie in the set $[0, 1]$. Thus, we may denote the function as $f : S \to S$, where $S = [0, 1]$. It is simply a formal way to say that for each input from the set S, the function produces a 'single' output which belongs to the same set S.

There exists a special element x^* in S such that when we feed this x^* as the input into the function, the output produced is also x^* (Figure 3). An element x^* in the set S is a fixed point of the function f (correspondence f^c) if $x^* = f(x^*)$ ($x^* \in f^c(x^*)$). Let us now think about the 'logical order' of the steps that we *must* use to talk about fixed point(s) of the above mentioned function.

[*Step 1*] We have to start with an underlying *structured set* – the one-dimensional Euclidean space $(\mathbb{R}, d_e)$, where $\mathbb{R}$ is the set of real numbers and d_e denotes the Euclidean metric on $\mathbb{R}$. It is only then that we can consider a subset of this structured set. It is important to note that although we did not mention it explicitly, the set $S = [0, 1]$ satisfies certain *properties*. It is non-empty, compact, and convex.

[*Step 2*] We can specify the *mapping* – $f : S \to S$ – only after we have specified the set S.

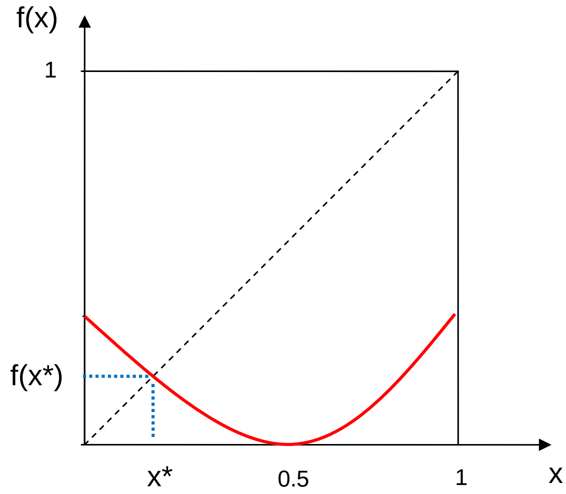

Figure 3: FIXED POINT OF A FUNCTION

Once again, although we did not mention it explicitly, this function satisfies certain properties. It is a continuous function. Let us now look at the general statement of *Brouwer's fixed point theorem*.

If

- S is a non-empty, compact, and convex subset of a finite dimensional Euclidean space; and,
- the mapping $f : S \to S$ is continuous,

then

- the mapping f will have at least one fixed point.

'A' describes the type and properties of the structured set and the type and properties of the mapping in Brouwer's fixed point theorem. The common pattern across fixed point theorems can be clarified by expressing the general form of almost all fixed point theorems.

If

- S is a subset of some *structured set* and satisfies certain *properties*; and,
- f is a certain type of *mapping* from S to S and satisfies certain *properties*,

then

- the *mapping* f will have at least one fixed point.

Table 8: Fixed Point Theorems

	Banach	**Kakutani**	**Glicksberg**
Structured Set	Metric Space	Euclidean Space	Normed Vector Space
Properties	Complete	Non-empty Compact Convex	Non-empty Compact Convex
Mapping	Function	Correspondence	Correspondence
Properties	Contractive	Convex valued Closed graphed	Convex valued Closed graphed

This general form highlights the usefulness of the classification of mathematical terms provided earlier. Table 8 lists the contents of 'A' in three other fixed point theorems. In general, we will find that 'A' refers to one or more of the four basic objects and their properties and 'B' specifies some additional properties of one or more of these basic objects in any result.

Understanding any result would require us to first understand the definitions of the properties associated with the basic objects. Although we do not go into a detailed description of these properties it should be clear that the definitions of these properties are also stated using the classification scheme we have described.

B. Arrow's impossibility theorem and related literature

We shall present the set up of the theorem in a stepwise manner and re-examine each step in the set up (see Appendix III for the formal set up). By its very nature, this process leads to a selective synthesis of the *questions* whose origins can be traced to Arrow's theorem .

[*Step 1*] $\mathbb{V}$ is a finite unstructured set containing $N \geq 2$ voters and $\mathbb{A}$ is a finite unstructured set containing $M \geq 3$ candidates.

Description: We assume the existence of an unstructured set $\mathbb{V}$ that contains a finite number of voters whose opinions will be used to determine the outcome of the election. $\mathbb{V}$ is indeed an unstructured set since the only thing we know is who are the irreducible elements of this set. Similarly, we assume the existence of an unstructured set $\mathbb{A}$ that contains a finite number of candidates who are standing in the election.

Questions: We could ask what if the set of voters is (un)countably infinite. However, it is hard to find interesting cases that should be modeled in this way. The case where the set of candidates is not finite captures features of several real world issues (Gaertner, 2006).

[*Step 2*] $\mathbb{L}$ is the finite structured set containing all linear orders (strict rankings) on $\mathbb{A}$. $\mathbb{L}^N$ is the N-fold cartesian product of the structured set $\mathbb{L}$.

Description: In this step we assume that each voter possesses a binary relation (the tool) that satisfies transitivity, completeness, and antisymmetry. This allows each voter to come up with a strict ranking over the set of candidates. The set $\mathbb{L}$ contains all the possible ways in which the candidates can be strictly ranked. If there are M candidates, then there would be $(M!)$ number of ways in which they could be strictly ranked.[10] If there are N voters, then we will receive N strict rankings as the *input* during the election. Since each voter may provide any of the strict rankings in the set $\mathbb{L}$ as her individual ranking, the set of possible inputs will be the N-fold cartesian product of set $\mathbb{L}$. Thus, the number of possible inputs will be $(M!)^N$.

Questions: Reflections on this step have led to a vast amount of research. We list some of the main questions that have motivated this research.

- Why should the voters be asked to send a complete (weak or strict) ranking of the candidates? What if they are asked to report only their most preferred candidate, or the subset of candidates they approve of? In general, these questions ask: What is the structure of the input provided by each voter in the social choice process?
- We implicitly assume intra-voter preferences to be ordinal and inter-voter preferences to be non-comparable when we ask voters to provide a ranking over the candidates. How can (and, what if) we allow for intra-voter preferences to be cardinal and inter-voter preferences to be comparable? If so, what should be the extent of comparability? Answers to these questions offer one way to understand what does it take to overcome the impossibility result (Sen, 1977).
- Will a voter have the *incentive* to truthfully reveal her ranking over the candidates? The literature on strategy proof mechanism design and implementation theory essentially springs from this question (Gibbard, 1973; Satterthwaite, 1975).

[*Step 3*] The function $f : \mathbb{L}^n \to \mathbb{L}$ denotes a Social Choice Rule.

Description: The function f can be thought of as a machine that uses a collection of N individual rankings as the input to produce one strict (social) ranking of the candidates as the output. Since the set $\mathbb{L}$ contains all possible ways in which the candidates can be strictly ranked, the task of f is to provide an output belonging to the set $\mathbb{L}$ for *each* of the $(M!)^N$ inputs belonging to the set $\mathbb{L}^N$. This is often referred to as the unrestricted domain (of voters' preferences) assumption.

[10] $M! = M \times (M-1) \ldots \times 2 \times 1$.

Questions: Is it reasonable to ask the Social Choice Rule to provide an output for every possible collective input? Are there situations where individual preferences have a certain structure so that we only need to search for social choice rules that satisfy our chosen properties on a restricted domain of preferences? The literature on single-peaked preferences provides an answer to this question. Clearly, as with the input, another question we could ask is: What happens if we want a weak-order over the set of candidates as the output? Or, just the winner of the election, a subset of the set of candidates, etc?

[*Step 4*] The Social Choice Rule (SCR) should satisfy the axioms of Weak Pareto (WP), Independence of Irrelevant Alternatives (IIA), and No-Dictatorship (ND).

Description: The axioms are used to impose restrictions on the structure of output given the features of the input. For example, suppose an input is such that all voters agree on the relative ranking of a certain pair of candidates. Axiom WP requires that the output corresponding to this input must have this same relative ranking of these two candidates. IIA asks us to look at two possible inputs. Suppose the relative ranking of a certain pair of candidates is the same in both these inputs by all the voters. Axiom IIA requires that the relative ranking of these two candidates in the outputs corresponding to these two inputs must be the same. Axiom ND essentially forbids the SCR from producing the social ranking by only using the individual ranking provided by one particular voter.

Questions: The impact of using the axiomatic approach to design social choice rules goes beyond social choice theory. We can definitely ask what happens if we replace the particular axioms used by Arrow with a different set of axioms. More importantly, it has led researchers to ask two extremely useful questions.

- Individual rankings (weak/linear orders) provided by several voters need to be aggregated to come up with a social ranking of the candidates in Arrow's framework. Are there other interesting real world phenomena where inputs from multiple sources need to be aggregated/ranked but the individual inputs are not rankings? Journals can be ranked according to their influence on scientific activity (Palacios-Huerta and Volij, 2004), theories can be ranked according to their predictive accuracy (Gilboa and Schmeidler, 2003), and researchers may be ranked according to their productivity (Woeginger, 2008). We may even aggregate the judgements of a group of people over a set of logical propositions (List and Petit, 2001; Dietrich, 2007). The input is not a ranking in any of these cases.

- Almost every aggregation problem requires us to specify axioms regarding how we want to convert inputs into outputs. What are the key *principles* that underlie the axioms (used by Arrow and in the subsequent literature)? Thomson (2000) provides an illuminating discussion of this question and identifies the few basic principles that form the conceptual basis behind the enormous variety of axioms that have been proposed in social choice theory and related fields.

Figure 4: The Common Procedure

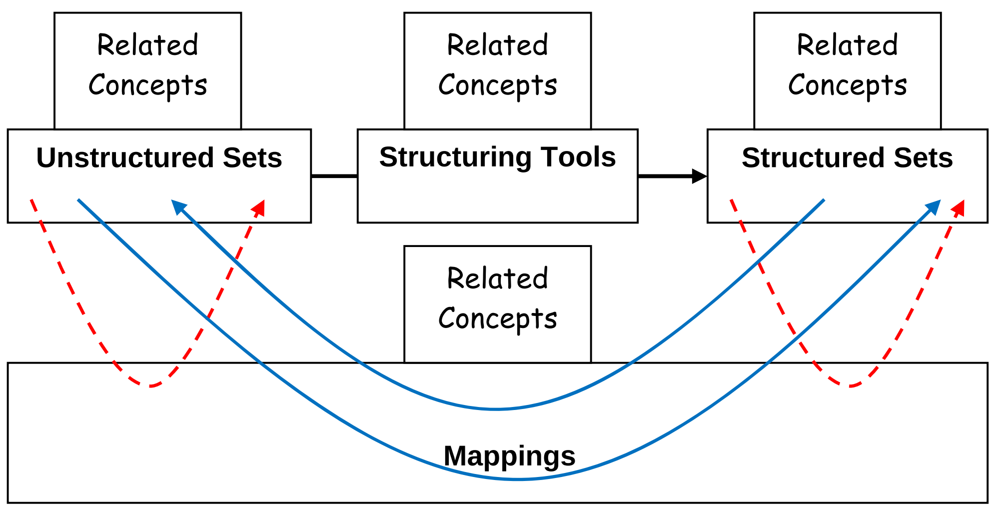

Figure 5: A classification of mathematical terms.

Unstructured sets and related concepts

We refer to a set of irreducible elements as an *unstructured set* if we know nothing more than the identity of the elements it contains. The notion of an unstructured set is the starting point of our discussion (please refer to Table 2 for the technical notation). We can talk about some notions related to sets in a meaningful way even if we know nothing more than the irreducible elements of the set(s) we are considering. Subset of a set, cardinality of a set, power set of a set, union and intersection of two sets, difference of two sets, disjoint subsets of a set, partition of a set, cartesian product of two sets, and n-fold cartesian product of a set are examples of such notions.

It is crucial to note that it takes certain axioms of set theory to assert the existence of, say, a power set of any given set or the union of two sets. Moreover, one may use different combinations of axioms to formulate a theory of sets. Zermelo-Fraenkel set theory with the axiom of choice (ZFC) is the dominant framework used by economists. Several researchers have questioned the reliance on ZFC to mathematically represent economic notions because of its non-constructivist features: the existence of some mathematical objects is assumed in ZFC even if one can not provide an algorithm to show how those objects may be constructed. The interested reader may refer to Vellupillai (2010) for a critical discussion of the drawbacks of ZFC and suggestions for a constructive framework for economic modeling.

Let $\mathbb{X} = \{x_1, x_2\}$ and $\mathbb{Y} = \{x_1, y_1, y_2\}$ be two sets.

- $x_2 \in \mathbb{X}$ and $y_1 \in \mathbb{Y}$.
- $\{x_1\} \subset \mathbb{X}$ (i.e., the set $\{x_1\}$ is a proper subset of $\mathbb{X}$).
- $\{x_1, y_1\} \subset \mathbb{Y}$ (i.e., the set $\{x_1, y_1\}$ is a proper subset of $\mathbb{Y}$).
- The set $\mathbb{X}$ is not equal to set $\mathbb{Y}$. In general, two sets are equal if they contain exactly the same irreducible elements. Note that this is an *axiom* of ZFC - the axiom of *extensionality.*
- $|\mathbb{X}| = 2$ and $|\mathbb{Y}| = 3$. The cardinality of a set may be finite, countably infinite, or uncountably infinite. A countably infinite set contains as many elements as contained in the set of natural numbers. Often, we use the term countable to refer to sets that either have finite or countably infinite number of elements. An uncountably infinite set is one that is not countable. The axiom of *infinity* is required to ensure that there exists a set with infinite elements.
- Power set of $\mathbb{X}$ is denoted by $P(\mathbb{X}) = \{\emptyset, \{x_1\}, \{x_2\}, \{x_1, x_2\}\}$. The axiom of *power set* is used to assert the existence of the power set of a given set.
- Power set of $\mathbb{Y} = P(\mathbb{Y}) = \{\emptyset, \{x_1\}, \{y_1\}, \{y_2\}, \{x_1, y_1\}, \{x_1, y_2\}, \{y_1, y_2\}, \{x_1, y_1, y_2\}\}$.
- In general, if $|\mathbb{X}| = N$, then $|P(\mathbb{X})| = 2^N$, where N is a natural number.
- Union of $\mathbb{X}$ and $\mathbb{Y} = \mathbb{X} \cup \mathbb{Y} = \{x_1, x_2, y_1, y_2\}$.

- Intersection of $\mathbb{X}$ and $\mathbb{Y} = \mathbb{X} \cap \mathbb{Y} = \{x_1\}$.
- Difference of $\mathbb{X}$ and $\mathbb{Y} = \mathbb{X} - \mathbb{Y} = \{x_2\}$. (Another notation is: $\mathbb{X} \backslash \mathbb{Y}$.)
- $\{x_1\}$ and $\{y_1\}$ are disjoint subsets of $\mathbb{Y}$ but do not constitute a partition of $\mathbb{Y}$.
- $\{y_1\}$ and $\{x_1, y_2\}$ are disjoint subsets of $\mathbb{Y}$ that constitute a partition of $\mathbb{Y}$.
- Cartesian Product of set $\mathbb{X}$ and set $\mathbb{Y}$ is denoted as $\mathbb{X} \times \mathbb{Y}$. Note that $\mathbb{X} \times \mathbb{Y}$ is itself a set. Similarly, the Cartesian Product of set $\mathbb{Y}$ and set $\mathbb{X}$ is denoted as $\mathbb{Y} \times \mathbb{X}$.
- $\mathbb{X} \times \mathbb{Y} = \{(x_1, x_1), (x_1, y_1), (x_1, y_2), (x_2, x_1), (x_2, y_1), (x_2, y_2)\}$. Each irreducible element of the set $\mathbb{X} \times \mathbb{Y}$ is an *ordered pair* where the first element in the pair comes from the set $\mathbb{X}$ and the second comes from the set $\mathbb{Y}$.
- $\mathbb{Y} \times \mathbb{X} = \{(x_1, x_1), (x_1, x_2), (y_1, x_1), (y_1, x_2), (y_2, x_1), (y_2, x_2)\}$. The irreducible elements of the sets $\mathbb{X} \times \mathbb{Y}$ and $\mathbb{Y} \times \mathbb{X}$ are different.
- The 2-fold cartesian product of set $\mathbb{X}$ with itself is denoted as $\mathbb{X} \times \mathbb{X}$ (or, as $\mathbb{X}^2$).
- $\mathbb{X}^2 = \{(x_1, x_1), (x_1, x_2), (x_2, x_1), (x_2, x_2)\}$. Each irreducible element of $\mathbb{X}^2$ is, once again, an *ordered pair*.[11] The term 'ordered' is used to stress that the element (x_1, x_2) should not be considered the same as the element (x_2, x_1).

[11] If we consider the 3-fold Cartesian product of $\mathbb{X}$ with itself, then each irreducible element of the resulting set ($\mathbb{X}^3$) will be an ordered triple.

Structured sets via multiple tools

Given an unstructured set $\mathbb{X}$, a *binary relator R defined over set* $\mathbb{X}$ is a tool that can be used to describe the relationship between two irreducible elements of $\mathbb{X}$. If x and y are irreducible elements of the unstructured set $\mathbb{X}$, then xRy is read as 'x is related to y.' It is crucial to note that the notion of a binary relator can be invoked only after we have specified the unstructured set.

Given an unstructured set $\mathbb{X}$, a *binary operator defined over set* $\mathbb{X}$ can be thought of as a machine that takes any pair of irreducible elements from $\mathbb{X}$ as the input and produces an irreducible element belonging to $\mathbb{X}$ as the output. Let $x \diamond y$ denote the output when we first feed x and then y into the binary operator denoted by $\diamond$. Once again, the notion of a binary operator can be introduced only after we have specified the set over which it will operate.

We have already looked at some structured sets that are obtained by using one binary relator or one binary operator as the structuring tool. Now we look at some structured sets that are obtained by using multiple tools.

[] The structured set $(\mathbb{X}, \oplus, \otimes)$ is a *Field* if

- it is an Abelian Group with respect to the binary operator $\oplus$;
- it is an almost-Abelian Group[12] with respect to the binary operator $\otimes$; and
- $\otimes$ distributes over $\oplus$ (i.e., $x \otimes (y \oplus z) = (x \otimes y) \oplus (x \otimes z)$, for any $x, y, z \in \mathbb{X}$).

[] The structured set $\mathbb{Q} \equiv (\mathbb{X}, \oplus, \otimes, R)$ is the set of *Rational numbers* if

- it is a field with respect to the two binary operators $\oplus$ and $\otimes$;
- the binary relator R is a linear order;
- R is compatible with the binary operators such that
  - xRy implies $(x \oplus z)R(y \oplus z)$, for any $x, y, z \in \mathbb{X}$; and,
  - $xRs_{\oplus}$ and $yRs_{\oplus}$ imply $(x \otimes y)Rs_{\oplus}$, for any $x, y \in \mathbb{X}$.

[] The structured set $\mathbb{R} \equiv (\mathbb{X}, \oplus, \otimes, R)$ is the set of *Real numbers* if

- it satisfies all the properties to be the set of Rational numbers; and,
- every non-empty $\mathbb{S} \subset \mathbb{X}$ with an upper bound in $\mathbb{X}$ has a supremum 'in $\mathbb{X}$.'

[12] $\tilde{x}_{\otimes}$ does not belong to $\mathbb{X}$ if $x = e_{\oplus}$.

[] Let S be any non-empty subset of $\mathbb{R}$. The set of

- *upper bounds* of S is $U(S) = \{u \in \mathbb{R} : u \geq s \ \forall \ s \in S\}$.
- *lower bounds* of S is $L(S) = \{l \in \mathbb{R} : l \leq s \ \forall \ s \in S\}$.

[] Let S be any non-empty subset of $\mathbb{R}$. Then

- the *supremum of S in* $\mathbb{R}$ is the unique smallest element in $U(S)$.
- the *infimum of S in* $\mathbb{R}$ is the unique largest element in $L(S)$.
- the supremum and infimum of S always exist.
- the supremum (infimum) is also referred to as the least upper bound (greatest lower bound).

[] Let S be any non-empty subset of $\mathbb{R}$. Then

- a *maximum of S in* $\mathbb{R}$ is an element $s^* \in S$ with $s^* \geq s \ \forall s \in S$.
- a *minimum of S in* $\mathbb{R}$ is an element $s_* \in S$ with $s_* \leq s \ \forall s \in S$.
- a maximim and/or minimum of S may not exist.

Metric Space

[] Let $\mathbb{X}$ be any unstructured nonempty set.

[] Let $d : \mathbb{X} \times \mathbb{X} \to \mathbb{R}$ be a *function* that assigns a real number to any pair of elements in the set $\mathbb{X}$. Intuitively, it takes a pair of elements in the set $\mathbb{X}$ as the input and provides the distance between them (some real number) as the output. We could also say that d takes an element from the set $\mathbb{X} \times \mathbb{X}$ as the input and produces a real number as the output.

[] The tuple $(\mathbb{X}, d)$ is called a *metric space* if the following three conditions hold.

- *Non-negativity*: $d(x, y) \ \geq \ 0 \ \forall x, y \in \mathbb{X}$, and $d(x, y) \ = \ 0$ iff $x = y$.
- *Symmetry*: $d(x, y) \ = \ d(y, x) \ \forall x, y \in \mathbb{X}$.
- *Triangular inequality*: $d(x, y) \ \leq \ d(x, z) + d(z, y) \ \forall x, y, z \in \mathbb{X}$.

[] The function d is called *a metric on the set* $\mathbb{X}$.

[] Unless necessary, we will denote the metric space as $\mathbb{X}$.

[] *ε-neighborhood of x in* $\mathbb{X}$: For any $x \in \mathbb{X}$ and $\varepsilon \in \mathbb{R}_{++}$, the *$\varepsilon$-neighborhood of x in X* is

- the set $N_{\varepsilon,\mathbb{X}}(x) = \{y \in \mathbb{X} : d(x, y) < \varepsilon\}$.

[] *Open subset of* $\mathbb{X}$: A subset S of $\mathbb{X}$ is said to be *open in* $\mathbb{X}$ if

- for each $x \in S$,
- $\exists \ \varepsilon \in \mathbb{R}_{++}$ such that,
- $N_{\varepsilon,\mathbb{X}} (x) \subseteq S$.

[] *Closed subset of* $\mathbb{X}$: A subset S of $\mathbb{X}$ is said to be *closed in* $\mathbb{X}$ if

- $\mathbb{X} \backslash S$ is open in $\mathbb{X}$.

[] $S \subset \mathbb{X}$ may be open, not open, closed, not closed, clopen, or neither open nor closed.

[] *Closure of a subset of* $\mathbb{X}$: Given any $S \subseteq X$, the closure of S in $\mathbb{X}$ is

- the smallest closed set $Cl_X(S)$ such that
- $S \subseteq Cl_X(S)$.

[] *Interior of a subset of* $\mathbb{X}$: Given any $S \subseteq X$, the interior of S in $\mathbb{X}$ is

- the largest open set $In_X(S)$ such that
- $In_X(S) \subseteq S$.

[] *Boundary of a subset of* $\mathbb{X}$: Given any $S \subseteq X$, the boundary of S in $\mathbb{X}$ is

- $Cl_X(S) - In_X(S)$.

[] *Convergent sequence in* $\mathbb{X}$: A sequence $\{x^m\} \in X^\infty$ converges to $x^* \in \mathbb{X}$ if

- for each $\epsilon > 0$
- $\exists$ a real number $M(\epsilon)$ such that
- $d(x^m, x^*) < \epsilon$
- $\forall m \geq M(\epsilon)$.

[] *Closed subset of* $\mathbb{X}$: A subset S is closed in $\mathbb{X}$ if, and only if,

- any sequence all of whose terms are in S
- converges to a point in S, if it converges at all.

[] *Bounded subset of* $\mathbb{X}$: A subset S is bounded in $\mathbb{X}$ if

- $\exists\ \epsilon > 0$ such that
- $S \subseteq N_{\epsilon,X}(x)$
- for some $x \in S$.

[] *Connected subset of* $\mathbb{X}$: A subset S is connected in $\mathbb{X}$ iff

- S can not be written as $S \equiv S_1 \cup S_2$
- where $S_1 \cap S_2 = \emptyset$
- and S_1, S_2 are open in S.

[] *Connected Metric Space* $\mathbb{X}$: $\mathbb{X}$ is connected iff the only clopen subsets of $\mathbb{X}$ are $\mathbb{X}$ and $\emptyset$.

[] *Dense subset of* $\mathbb{X}$: A subset Y is said to be dense in $\mathbb{X}$ if

- $Cl_X(Y) = X$.

[] *Separable Metric Space* $\mathbb{X}$: $\mathbb{X}$ is separable if

- $\exists\ Y \subseteq X$ such that
- $Cl_X(Y) = X$
- and Y is countable.

[] *Cover of* $S \subseteq \mathbb{X}$: A collection $\mathcal{O}$ of subsets of $\mathbb{X}$ is a cover of S in $\mathbb{X}$ if

- $S \subseteq \cup\mathcal{O}$.

[] *Open cover of* $S \subseteq \mathbb{X}$: A collection $\mathcal{O}$ of subsets of $\mathbb{X}$ is an open cover of S in $\mathbb{X}$ if

- $S \subseteq \cup\mathcal{O}$
- and each member of $\mathcal{O}$
- is open in $\mathbb{X}$.

[] *Compact subset of* $\mathbb{X}$: A subset S is compact in $\mathbb{X}$ if

- every finite and infinite open cover of S in $\mathbb{X}$
- has a finite subset
- that also covers S.

[] *Totally bounded subset of* $\mathbb{X}$: A subset S is totally bounded in $\mathbb{X}$ if

- for any $\epsilon > 0$
- $\exists$ a finite subset T of S such that
- $S \subseteq \cup\{N_{\epsilon,X}(x) :\ x \in T\}$.

[] *Cauchy sequence in* $\mathbb{X}$: A sequence $\{x^m\} \in X^\infty$ is said to be cauchy if

- for each $\epsilon > 0$
- $\exists$ a real number $M(\epsilon)$ such that
- $d(x^k, x^l) < \epsilon$
- $\forall k, l \geq M(\epsilon)$.

[] *Complete Metric Space* $\mathbb{X}$: A metric space is complete if every cauchy sequence in $\mathbb{X}$

- converges to a point in $\mathbb{X}$.

[] A *function* $f : \mathbb{X} \to \mathbb{Y}$ takes elements from $\mathbb{X}$ as the input and provides elements from $\mathbb{Y}$ as the output. It is important to note that

- f produces only one element from $\mathbb{Y}$ as the output for any $x \in \mathbb{X}$.
- It must produce an output for every $x \in \mathbb{X}$ otherwise it does not qualify as a function.
- $f(x) \in \mathbb{Y}$ is called the image of $x \in \mathbb{X}$.
- $x \in \mathbb{X}$ is called the pre-image of $f(x) \in \mathbb{Y}$.
- We may thus say that for each $x \in \mathbb{X}$ the function produces a *unique* image $f(x) \in \mathbb{Y}$.

[] A self-map $f : (\mathbb{X}, d_X) \to (\mathbb{X}, d_X)$ is a *contraction* if

- $\exists\ K \in (0, 1)$ such that
- $d(f(x), f(y)) \leq K d(x, y)$
- $\forall x, y \in \mathbb{X}$.

[] Given a non-empty set $\mathbb{X}$, let $\overline{P(\mathbb{X})}$ denote the collection of all non-empty subsets of $\mathbb{X}$.

[] A *correspondence* $f^c : \mathbb{X} \to \overline{P(\mathbb{Y})}$ takes elements from $\mathbb{X}$ as the input and provides elements from $\overline{P(\mathbb{Y})}$ as the output. It is important to note that

- f^c produces only one element from $\overline{P(\mathbb{Y})}$ as the output for any $x \in \mathbb{X}$; but one element of $\overline{P(\mathbb{Y})}$ may contain one or more elements of $\mathbb{Y}$.
- It must produce an output for every $x \in \mathbb{X}$ otherwise it does not qualify as a correspondence.
- $f^c(x) \in \overline{P(\mathbb{Y})}$ is called the image of $x \in \mathbb{X}$.
- $x \in \mathbb{X}$ is called the pre-image of $f^c(x) \in \overline{P(\mathbb{Y})}$.
- We will denote $f^c : \mathbb{X} \to \overline{P(\mathbb{Y})}$ as $f^c : \mathbb{X} \rightrightarrows \mathbb{Y}$ throughout this section.

Functions over Metric Spaces

[] Given a function $f : (\mathbb{X}, d_X) \to (\mathbb{Y}, d_Y)$, let

- O_Y denote any open set in $(\mathbb{Y}, d_Y)$
- $O_Y(y)$ denote any open set in $(\mathbb{Y}, d_Y)$ that contains $y \in \mathbb{Y}$.
- $f^{-1}(O_Y)$ denote the subset $\{x \in \mathbb{X} : f(x) \in O_Y\}$ (the pre-image of O_Y).
- $f^{-1}(O_Y(y))$ denote the subset $\{x \in \mathbb{X} : f(x) \in O_Y(y)\}$ (the pre-image of $O_Y(y)$).

[] A function $f : (\mathbb{X}, d_X) \to (\mathbb{Y}, d_Y)$ is

- *discontinuous at* $x^* \in \mathbb{X}$ if $f^{-1}(O_Y(f(x^*))$ is not open for at least one $O_Y(f(x^*))$.
- *continuous at* $x^* \in \mathbb{X}$ if $f^{-1}(O_Y(f(x^*))$ is open for every $O_Y(f(x^*))$.

[] A function $f : (\mathbb{X}, d_X) \to (\mathbb{Y}, d_Y)$ is

- *discontinuous over* $\mathbb{X}$ if $f^{-1}(O_Y)$ is not open for at least one O_Y.
- *continuous over* $\mathbb{X}$ if $f^{-1}(O_Y)$ is open for every O_Y.

[] Given a function $f : (\mathbb{X}, d_X) \to (\mathbb{R}, d_e)$

- the *strict lower contour set* of $\alpha \in \mathbb{R}$ is $L_X(\alpha) = \{x \in \mathbb{X} : f(x) < \alpha\}$.
- the *strict upper contour set* of $\alpha \in \mathbb{R}$ is $U_X(\alpha) = \{x \in \mathbb{X} : f(x) > \alpha\}$.

[] A function $f : (\mathbb{X}, d_X) \to (\mathbb{R}, d_e)$ is

- USC over $\mathbb{X}$ if $L_X(\alpha)$ is open for every $\alpha \in \mathbb{R}$.
- LSC over $\mathbb{X}$ if $U_X(\alpha)$ is open for every $\alpha \in \mathbb{R}$.

[] Given a correspondence $f^c : (\mathbb{X}, d_X) \rightrightarrows (\mathbb{Y}, d_Y)$, let

- O_Y denote any open set in $(\mathbb{Y}, d_Y)$
- $O_Y(S)$ denote any open set in $(\mathbb{Y}, d_Y)$ that contains $S \subseteq \mathbb{Y}$.
- the *upper inverse image* of O_Y be $f^{-1c}(O_Y) = \{x \in \mathbb{X} : f^c(x) \subseteq O_Y\}$.
- the *lower inverse image* of O_Y be $f_{-1c}(O_Y) = \{x \in \mathbb{X} : f^c(x) \cap O_Y \neq \emptyset\}$.

[] A correspondence $f^c : (\mathbb{X}, d_X) \rightrightarrows (\mathbb{Y}, d_Y)$ is

- *not* UHC at $x^* \in \mathbb{X}$ if $f^{-1c}(O_Y(f^c(x^*))$ is not open for at least one $O_Y(f^c(x^*))$.
- *not* LHC at $x^* \in \mathbb{X}$ if $f_{-1c}(O_Y(f^c(x^*))$ is not open for at least one $O_Y(f^c(x^*))$.

[] A correspondence $f^c : (\mathbb{X}, d_X) \rightrightarrows (\mathbb{Y}, d_Y)$ is

- UHC at $x^* \in \mathbb{X}$ if $f^{-1c}(O_Y(f^c(x^*))$ is open for every $O_Y(f^c(x^*))$.
- LHC at $x^* \in \mathbb{X}$ if $f_{-1c}(O_Y(f^c(x^*))$ is open for every $O_Y(f^c(x^*))$.

[] A correspondence $f^c : (\mathbb{X}, d_X) \rightrightarrows (\mathbb{Y}, d_Y)$ is

- UHC iff $f^{-1c}(O_Y)$ is open for every O_Y.
- LHC iff $f_{-1c}(O_Y)$ is open for every O_Y.

[] A correspondence $f^c : (\mathbb{X}, d_X) \rightrightarrows (\mathbb{Y}, d_Y)$ is *compact-valued* if

- $f^c(x)$ is a compact subset of $\mathbb{Y}$ for each $x \in \mathbb{X}$.

[] A correspondence $f^c : (\mathbb{X}, d_X) \rightrightarrows (\mathbb{Y}, d_Y)$ is *closed-valued* if

- $f^c(x)$ is a closed subset of $\mathbb{Y}$ for each $x \in \mathbb{X}$.

[] A correspondence $f^c : (\mathbb{X}, d_X) \rightrightarrows (\mathbb{R}^n, d_e)$ is *convex-valued* if

- the set $f^c(x)$ is convex for each $x \in \mathbb{X}$.

[] A correspondence $f^c : (\mathbb{X}, d_X) \rightrightarrows (\mathbb{Y}, d_Y)$ is *closed at* $x \in \mathbb{X}$ if

- for any convergent sequences $(x^m) \in \mathbb{X}^\infty$ and $(y^m) \in \mathbb{Y}^\infty$ with $x^m \to x$ and $y^m \to y$
- $y \in f^c(x)$ if $y^m \in f^c(x^m)$ for each $m \geq 1$.

[] A correspondence $f^c : (\mathbb{X}, d_X) \rightrightarrows (\mathbb{Y}, d_Y)$ has a *closed-graph* $x \in \mathbb{X}$ if

- it is closed at every $x \in \mathbb{X}$.

Vector Space

[] *Vector Space*: Consider a set $\mathbb{X}$ whose irreducible elements may be denoted as x, y, z. In order to obtain the structured set called a vector space, we need to use two binary operators as the tools. Let us denote these tools as $\oplus$ and $\odot$. It will be helpful to think of $\oplus$ as a tool that allows us to add any two elements of $\mathbb{X}$ and $\odot$ as a tool that can be used to multiply a real number with any element of $\mathbb{X}$. The structured set $\mathbb{V} \equiv (\mathbb{X}, \oplus, \odot)$ is referred to as a 'Vector Space using real numbers' and every element of $\mathbb{X}$ is referred to as a 'vector' if

- $(\mathbb{X}, \oplus)$ is an Abelian-group (which implies that $\mathbb{V}$ contains a special element $s_{\oplus}$ which plays the same role as 0 in the case of real numbers);
- $\alpha \odot x \in \mathbb{X}$, for any $\alpha \in \mathbb{R}$ and any $x \in \mathbb{X}$; and
- for any $x, y \in \mathbb{X}$ and any $\alpha, \beta \in \mathbb{R}$,
  - $\alpha \odot (\beta \odot x) = (\alpha \odot \beta) \odot x$;
  - $1 \odot x = x$;
  - $(\alpha \oplus \beta) \odot x = \alpha \odot x \oplus \beta \odot x$; and,
  - $\alpha \odot (x \oplus y) = \alpha \odot x \oplus \alpha \odot y$.

[] The Vector Space used most often in economics is one where $\mathbb{X} = \mathbb{R}^n$.

[] What is $\mathbb{R}^n$?
[] It is crucial to note that

- we don't know the 'distance between any pair of elements' in $\mathbb{R}^n$;
- we don't know the 'distance from origin' for any element of $\mathbb{R}^n$;
- there is no order structure over the elements of $\mathbb{R}^n$ if $n > 1$;
- there is no notion of a line or line segment in $\mathbb{R}^n$ if $n > 1$.

[] Similarly,

- we don't know the 'distance between any pair of elements' in $\mathbb{V}$;
- we don't know the 'distance from origin' for any element of $\mathbb{V}$;
- there is no order structure over the elements of $\mathbb{V}$; but,
- there is a well defined notion of a line or line segment in $\mathbb{V}$.

[] Consider the following regression model.

$$\mathbf{y} \;=\; \beta_0 \cdot \mathbf{1_N} \;+\; \beta_1 \cdot \mathbf{x} \;+\; \beta_2 \cdot \mathbf{z} \;+\; \mathbf{u},$$

where $\mathbf{y}$, $\mathbf{x}$, $\mathbf{z}$, and $\mathbf{u}$ are N-dimensional column vectors. Irrespective of the economic meaning of the variables, when we actually run the regression each entry in these column vectors is a real number. Thus, these vectors are irreducible elements of the set $\mathbb{R}^N$. It should now be clear that when we run this regression, we are dealing with the vector space

$$(\mathbb{X}, \oplus, \odot) \quad \equiv \quad (\mathbb{R}^N, +, \times),$$

where the binary operator $+$ is used to add two elements of $\mathbb{R}^N$ and the binary operator $\times$ is used to multiply a real number with an element of $\mathbb{R}^N$.

A Vector space is also referred to as a *linear space*. The reason being that we can define the notion of a line segment in a vector space. Given any $x, y \in \mathbb{X}$, if the structured set $(\mathbb{X}, \oplus, \odot)$ is a vector space, then the line segment joining x and y is the set

$$\mathbb{L}_{xy} \;=\; \{\alpha \;\otimes\; x \;\oplus\; (1-\alpha) \;\otimes\; y \mid \alpha \in [0,1]\}.$$

If we can define the notion of a line segment in a structured set, then we have the means to say whether or not a subset of this set is *convex*. A subset of a structured set is convex if the line joining any two points in the subset lies in the subset. Several fundamental results in economic theory rely heavily on the notion of convexity of a set (Ok, 2006, Chapter H).

[] *Normed Vector Space*: Suppose we have a vector space $\mathbb{V}$. We can use a variety of tools to provide some additional structure to the elements of $\mathbb{V}$. We just saw that there is no notion of 'distance from origin' for any element of $\mathbb{V}$. The concept of a 'norm' is used to provide some additional structure to $\mathbb{V}$ by ensuring that we can meaningfully talk about the 'distance from origin.' The structured set $(\mathbb{V}, \eta)$ is a normed vector space if the function $\eta : \mathbb{V} \to \mathbb{R}$ satisfies the following three conditions for any $x, y \in \mathbb{V}$.

- *Non-negativity*: $\eta(x) \geq 0$, and $\eta(x) = 0$ iff $x = \mathbf{0}$.
- *Homogeneity*: $\eta(\alpha x) = \alpha\eta(x)$, for any $\alpha \in \mathbb{R}_{++}$.
- *Triangular inequality*: $\eta(x+y) \leq \eta(x) + \eta(y)$.

[] If η is a norm over $\mathbb{V}$, then defining $d(x,y) = \eta(x-y)$ gives us a metric over $\mathbb{V}$.[13] If you give me a norm over $\mathbb{V}$, I can use it to define a metric over $\mathbb{V}$. On the other hand, if I give you a metric over $\mathbb{V}$, you may or may not be able to use it to define a norm over $\mathbb{V}$. In summary,

- every norm over $\mathbb{V}$ generates a metric;
- there are metrics over $\mathbb{V}$ that do not generate a norm over $\mathbb{V}$; and thus,

[13] Note that $(x-y) \in \mathbb{V}$ if $x, y \in \mathbb{V}$, by definition. This is why it makes sense to define the metric using the norm.

- every normed vector space qualifies as a metric space, but some metric spaces do not qualify as a normed vector space.

[] A *Banach Space* is a *complete* normed vector space. The norm induces a metric, which is turn allows us to talk about the convergence of cauchy sequences, and hence the completeness of a normed vector space.

Topological Space

Suppose $\mathbb{X} = \{a, b, c, d\}$ is some set and consider the following sets.

- $\mathbb{T}_1 = \{\emptyset, \{a\}, \{b, c, d\}, \{a, b, c, d\}\}$.
- $\mathbb{T}_2 = \{\{a\}, \{a, b, c, d\}\}$.
- $\mathbb{T}_3 = \{\emptyset, \{a\}, \{b\}, \{a, b, c, d\}\}$.
- $\mathbb{T}_4 = \{\emptyset, \{a, b\}, \{b, c, d\}, \{a, b, c, d\}\}$.
- $\mathbb{T}_5 = \{\emptyset, \{a, b, c, d\}\}$.

The irreducible elements of these four sets are subsets of the initial set we started with. Now consider the following three properties.

- The empty set and the set $\mathbb{X}$ must be irreducible elements of the set $\mathbb{T}_i$.
- The union of any number of irreducible elements of the set $\mathbb{T}_i$ must also be an irreducible element of $\mathbb{T}_i$.
- The intersection of any finite number of irreducible elements of the set $\mathbb{T}_i$ must also be an irreducible element of $\mathbb{T}_i$.

It can be easily verified that the set $\mathbb{T}_1$ and $\mathbb{T}_5$ satisfy each of these three properties, but the sets $\mathbb{T}_2, \mathbb{T}_3,$ and $\mathbb{T}_4$ do not satisfy at least one of the three properties. The sets $\mathbb{T}_1$ and $\mathbb{T}_5$ are referred to as *topologies on* $\mathbb{X}$. The structured sets $(\mathbb{X}, \mathbb{T}_1)$ and $(\mathbb{X}, \mathbb{T}_5)$ are termed *topological spaces*. We may thus say that in order to formulate a topological space we start with some set and the tool we use to add structure to it is a collection of subsets of the starting set.

The elements of $\mathbb{T}_1$ ($\mathbb{T}_5$) are referred to as *open sets* in the associated topological space. Thus, the subset $\{b, c, d\}$ is an open set in the topological space $(\mathbb{X}, \mathbb{T}_1)$ but it is not an open set in the topological space $(\mathbb{X}, \mathbb{T}_5)$. A quick summary of the main concepts follows.

[] Consider any non-empty set X. Let T be a collection of subsets of X. The structured set (X, T) is called a topological space if

- $\emptyset, X \in T$;
- Arbitrary unions and finite intersections of elements of T are also in T.

[] If (X, T) is a topological space, then

- T is called a topology on X; and,
- Subsets of X that are elements of T are referred to as *open* sets in (X, T).
- Subsets of X whose complements are elements of T are referred to as *closed* sets.

[] (X, T) is a Hausdorff topological space if

- for each distinct pair of elements $x, y \in X$
- $\exists\ t_x, t_y \in T$ such that
- $x \in t_x$, $y \in t_y$, and $t_x \cap t_y = \emptyset$.

[] $B \subset T$ is a basis for T if every element of T can be expressed as a union of members of B.

[] The smallest closed subset of (X, T) that contains the open subset A is referred to as the closure of A.

[] Any subset A of X is a dense subset of (X, T) if the closure of A equals X.

[] $N \subset X$ is a neighborhood of $x \in N$ if there exists an open set A such that $p \in U \subseteq N$.

[] (X, T) is said to be connected if the only clopen subsets are $\emptyset$ and X.

[] $f : (X, T_X) \to (Y, T_Y)$ is continuous if the pre-image of open subsets of (Y, T_Y) is open in (X, T_X).

[] (X, T_X) and (Y, T_Y) are said to be homeomorphic if there exists an $f : (X, T_X) \to (Y, T_Y)$ such that (i) f is one-to-one and onto, and (ii) both f and f^{-1} are continuous.

[] $A \subseteq X$ is said to be a compact if every open cover of A has a finite subcover.

[] Consider any non-empty set X and its power set $\mathbb{P}(X)$. $\mathcal{A} \subseteq \mathbb{P}(X)$ is a σ-algebra over X if

- $X \in \mathcal{A}$;
- $A \in \mathcal{A} \;\Rightarrow\; A^c \in \mathcal{A}$; and,
- Countable union of elements in $\mathcal{A}$ is also an element of $\mathcal{A}$.

[] The structured set $(X, \mathcal{A})$ is a measurable space if $\mathcal{A}$ is a sigma algebra. The elements of $\mathcal{A}$ are called measurable sets.

[] $f : (X, \mathcal{A}_X) \to (Y, \mathcal{A}_Y)$ is a measurable function if the pre-image of a measurable set in $(Y, \mathcal{A}_Y)$ is a measurable set in $(X, \mathcal{A})$.

[] The structured set $(X, \mathcal{A}, \mu)$ is a measure space if $\mathcal{A}$ is a sigma algebra and the function $\mu : \mathcal{A} \to \mathbb{R}_+ \cup \infty$ is such that

- $\mu(\emptyset) = 0$; and,
- $\mu(\cup A_i) = \sum \mu(A_i)$, for any countable number of pairwise disjoint elements of $\mathcal{A}$.

[] The structured set $(X, \mathcal{A}, P)$ is a probability space if

- $\mathcal{A}$ is a sigma algebra and P is a (probability) measure with $P(X) = 1$.
- X is the state space and elements of $\mathcal{A}$ are events.

[] Given a topological space (X, T), the collection $\mathcal{B}_t$ is a Borel-algebra if

- the set $\mathcal{B}_t$ is the unique smallest σ-algebra such that $\mathcal{B}_t \supseteq T$.

[] $f : (X, T_X) \to (Y, T_Y)$ is said to be Borel-measurable function if it is measurable once X and Y are equipped with their respective Borel-algebras.

[] Given $X \equiv (\mathbb{R}^n, d_e)$ and a Borel-algebra $\mathcal{B}_e$ over X, the (unique) Lebesgue measure $\mu_l : \mathcal{B}_e \to \mathbb{R}_+ \cup \infty$ is such that for any $B \in \mathcal{B}_e$,

- $\mu_l(B) \;=\; sup\{\mu(C) : C \subseteq B \text{ is compact}\} \;=\; inf\{\mu(O) : O \supseteq B \text{ is open}\}$.

[] $((\mathbb{R}^n, d_e), \mathcal{B}_e, \mu_l)$ is called a Lebesgue-measurable space.

Individual Choice Theory: The Basics

The Basic Set Up

[1] $\mathbb{X}$ is a non-empty finite (or, countably infinite) set of alternatives.
[2] R is a binary relation over the elements of $\mathbb{X}$ that can be used to provide some additional structure to the elements of set $\mathbb{X}$.
[3] For any pair of alternatives $x, y \in \mathbb{X}$, $[xRy]$ means [x *is at least as good as* y].
[4] Some useful properties of R are formally listed in the following Table.

Table 9: Properties of Binary Relations

R **is a**	**If**
Reflexive binary relation	$xRx, \ \forall x \in \mathbb{X}$
Transitive binary relation	xRy and yRz imply xRz, $\forall x, y, z \in \mathbb{X}$
Antisymmetric binary relation	xRy and yRx imply $x = y$, $\forall x, y \in \mathbb{X}$
Complete binary relation	xRy or yRx, $\forall x, y \in \mathbb{X}$
Symmetric binary relation	xRy implies yRx, $\forall x, y \in \mathbb{X}$
Asymmetric binary relation	xRy implies $\neg\ (yRx)$, $\forall x, y \in \mathbb{X}$

[5] A binary relation R can often be decomposed into a symmetric part and an asymmetric part. For example, suppose the binary relation R stands for 'at least as good as.' Then, we can define (i) the *symmetric* binary relation I_R derived from R to mean 'as good as', and (ii) the *asymmetric* binary relation P_R derived from R to mean 'strictly better than.' Note that:

- xRy implies $[xI_Ry$ or $xP_Ry]$
- xI_Ry implies $[xRy$ and $yRx]$ (i.e., I_R is symmetric)
- xP_Ry implies $[xRy$ and $\neg\ (yRx)]$ (i.e., P_R is asymmetric)

[6] We obtain different types of structured sets depending upon the properties of R. Some useful types of structured sets are listed in the following Table.

Table 10: Structured sets using binary relations as the structuring tool

$(\mathbb{X}, R)$ **is a**	**If** R **is**
Pre-ordered set	Reflexive, Transitive
Weakly-ordered set	Reflexive, Transitive, Complete
Partially-ordered set	Reflexive, Transitive, Antisymmetric
Linearly-ordered set	Reflexive, Transitive, Antisymmetric, Complete

[7] We will refer to R as the *preference relation* of the decision maker.

[8] *Rational Preference Relation*: If the decision maker can use his R to structure the elements of $\mathbb{X}$ into a weakly-ordered set or a linearly ordered set, then we *say* that the individual has rational preferences over the elements of the set $\mathbb{X}$.

[9] It is important to note that the word rational is a essentially a convenient (and, somewhat reasonable) label.

[10] Let $u : \mathbb{X} \to \mathbb{R}$ be some function that takes an element from the set $\mathbb{X}$ as the input and provides a real number as the output.

[11] We *say* that *a function* $u : \mathbb{X} \to \mathbb{R}$ *represents* the preference relation R if

- for any pair of alternatives $x, y \in \mathbb{X}$,
- xRy iff $u(x) \geq u(y)$.

[12] If we can find *at least one* function such that

- xRy iff $u(x) \geq u(y)$ for any pair of alternatives $x, y \in \mathbb{X}$,
- then we refer to the function u as a utility function,
- and *say* that *the preference relation R admits a utility function representation*.

[13] The concept of a utility function is simply a *re-formatting of the information* contained in the preference relation into a more readily usable form. In fact, one can do (almost) everything in economics without ever introducing the notion of a utility function. We often deal directly with utility functions instead of preference relations because it usually makes the math relatively easy.

[14] Some of the most basic questions of individual choice theory are:

- Can every preference relation be represented by a function?
- What properties must be satisfied by the preference relation so that it can definitely be represented by a function?
- If the preference relation can be represented by a function, then is this function unique?
- If a preference relation can be represented by many functions, and we know one of these functions, then can we find all the remaining such functions using only the one we know?

Utility Representation Theorem: When $|\mathbb{X}|$ is Uncountable

---

The Basic Set Up

[1] $\mathbb{X} \subset \mathbb{R}^L$ is an uncountable set of alternatives in the L-dimensional Euclidean Space.
[2] R is a binary preference relation over the elements of $\mathbb{X}$.
[3] For any pair of alternatives $x, y \in \mathbb{X}$, $[xRy]$ means [*x is at least as good as y*].
[4] Let $u : \mathbb{X} \to \mathbb{R}$ be some function that takes an element from the set $\mathbb{X}$ as the input and provides a real number as the output.

The Axioms

[W] *R is a rational binary preference relation*

- R is a weak-order.

[C] *R is a continuous binary preference relation*

- for any $x \in \mathbb{X}$,
- the sets $B(x) = \{y \in \mathbb{X} : \; yRx\}$ and $W(x) = \{y \in \mathbb{X} : \; xRy\}$
- are closed subsets of $\mathbb{X}$.

UR-Theorem 2: *Any continuous weak-order defined over a non-empty uncountable set $\mathbb{X} \subset \mathbb{R}^L$ admits representation by a continuous utility function; and, the utility function is unique up to strictly increasing transformations.*

Individual Choice Under Risk

The Basic Set Up

[1] Let $\mathbb{N}$ denote the set $\{1, 2, \ldots, n\}$ with cardinality N.
[2] $\mathbb{X}$ is a non-empty finite set of *outcomes* with $|\mathbb{X}| = N$. Let $\mathbb{X} = \{x_1, x_2, \ldots, x_k, \ldots, x_n\}$.
[3] Let $\mathcal{L}$ denote the set of all lotteries that can be constructed by using the set $\mathbb{X}$.
[4] Any *lottery* $l \in \mathcal{L}$ specifies the probabilities $\{p_k(l)\}_{k\in\mathbb{N}}$ with which the agent will obtain the outcome $\{x_k\}_{k\in\mathbb{N}}$. For any *lottery* $l \in \mathcal{L}$

- $p_k(l) \in [0, 1]\ \forall k \in \mathbb{N}$ and $\sum_{k\in\mathbb{N}} p_k(l) = 1$.

[5] What is the cardinality of the set $\mathcal{L}$? What does it look like?
[6] We are interested in understanding how an agent will make a choice when faced with a subset of *alternatives* from the set $\mathcal{L}$.
[7] R is a binary preference relation over the elements of $\mathcal{L}$. We shall denote the symmetric and asymmetric parts of R by I and P, respectively. Note that R is defined over an uncountable set.
[8] For any pair of lotteries $l, \tilde{l} \in \mathcal{L}$, $[l\ R\ \tilde{l}]$ means [*lottery l is at least as good as lottery $\tilde{l}$*].
[9] Given the set of outcomes $\mathbb{X}$, the *expected value* of any lottery $l \in \mathcal{L}$ is *defined* as

$$V(l|\mathbb{X}) \ = \ \sum_{k\in\mathbb{N}} (p_k(l) \cdot x_k).$$

[10] Let $u : \mathbb{X} \to \mathbb{R}$ be a utility function that represents preferences over the elements of $\mathbb{X}$. The *expected utility* of any lottery $l \in \mathcal{L}$ is *defined* as

$$U(l|\mathbb{X}) \ = \ \sum_{k\in\mathbb{N}} (p_k(l) \cdot u(x_k)).$$

[11] The set $\{\hat{l} = (\alpha \cdot l + (1-\alpha) \cdot \tilde{l}) : \alpha \in [0, 1]\}$ contains all the *compound* lotteries (convex combinations) that can be constructed using any pair of lotteries $l, \tilde{l} \in \mathcal{L}$.

The Axioms

[W] <u>*R is a rational binary preference relation*</u>

- R is a complete pre-order. (What does this mean for the preferences over $\mathbb{X}$?)

[C] <u>*R is a continuous binary preference relation*</u>

- For any three lotteries $l, \tilde{l}, \hat{l} \in \mathcal{L}$,
- if $l\ P\ \tilde{l}\ P\ \hat{l}$,
- then there exist $\alpha, \beta \in (0, 1)$ such that,
- $(\alpha \cdot l + (1-\alpha) \cdot \hat{l})\ P\ \tilde{l}\ P\ (\beta \cdot l + (1-\beta) \cdot \hat{l})$.

[IIA] <u>*R is an IIA binary preference relation*</u>

- For any $l, \tilde{l}, \hat{l} \in \mathcal{L}$ and any $\alpha \in (0,1)$,
- $[l \; R \; \tilde{l}]$ *iff* $[(\alpha \cdot l + (1-\alpha) \cdot \hat{l}) \; R \; (\alpha \cdot \tilde{l} + (1-\alpha) \cdot \hat{l})]$.

<u>*Interpretation of IIA*</u>

The IIA axiom formalizes the notion of *dynamic consistency.* It assumes that the agent would make the same choice in each of the following four cases.

[**Case 1**] The agent has to choose between the lotteries l and $\tilde{l}$.
[**Case 2**] The experimenter first chooses whether or not the agent will get a chance to make a choice. With probability $(1-\alpha)$ the agent will not get a chance to choose, and will be awarded the lottery $\hat{l}$; with the remaining probability α the agent gets to choose, and is asked to report his choice between the lotteries l and $\tilde{l}$.
[**Case 3**] Same as (ii) except that the agent is asked to report her choice between l and $\tilde{l}$ before knowing whether he will actually get the chance to choose between the lotteries l and $\tilde{l}$.
[**Case 4**] The agent is asked to choose between $(\alpha \cdot l + (1-\alpha) \cdot \hat{l})$ and $(\alpha \cdot \tilde{l} + (1-\alpha) \cdot \hat{l})$.

<u>*Experimental Procedure to test IIA*</u>

- The participants in the experiment will be presented each of these four scenarios and their choices will be observed.
- The fraction of participants who do not report the same answer for each of the four questions can be easily calculated.
- This procedure will be repeated for several combinations of the three lotteries denoted by $l, \tilde{l}$, and $\hat{l}$.
- If a significant fraction of participants report different answers to the four questions in all the repetitions of the experiment, then we can conclude that the Independence axiom does not provide an accurate description of individual decision making in the presence of risk.

VNM Theorem: The binary preference relation R satisfies W, C, and IIA

- *iff* there exists a function $u : \mathbb{X} \to \mathbb{R}$ such that,
  - for every pair of lotteries $l, \tilde{l} \in \mathcal{L}$,
  - $[l \; R \; \tilde{l}]$ iff $[U(l|\mathbb{X}) \;\geq\; U(\tilde{l}|\mathbb{X})]$.
- The utility function $u(\cdot)$ is unique up to positive linear transformations.

Subjective Probabilities I: De Finnetti's Theorem

The Basic Set Up

[1] Let $\mathbb{S} = \{s_1, s_2, \ldots, s_n\}$ denote the set of *states* with cardinality n.
[2] Let $\mathbb{X}$ be the vector space over $\mathbb{R}^n$. Any $x \in \mathbb{X}$ will be denoted as $(x_1, \ldots, x_n)$.
[3] The function $f : \mathbb{S} \to \mathbb{X}$ is a *bet* which specifies the amount of monetary gain or loss to the decision maker for each possible state that may be realized. Thus, the bet f essentially means the vector $(f(s_1), \ldots, f(s_n))$. Let $\mathcal{F}$ denote the set of all possible bets.
[4] We are interested in the decision makers ranking over the elements of $\mathcal{F}$.
[5] R is a binary preference relation over the elements of the uncountable set $\mathcal{F}$. We shall denote the symmetric and asymmetric parts of R by I and P, respectively.
[6] For any pair of bets $f, g \in \mathcal{F}$, $[fRg]$ means [*bet f is at least as good as bet g*].
[7] Note that the $[fRg]$ amounts to saying that [*the vector $(f(s_1), \ldots, f(s_n))$ is at least as good as the vector $(g(s_1), \ldots, g(s_n))$*].

The Axioms

[W] *R is a rational binary preference relation*

- R is a weak order defined over $\mathcal{F}$. (Or, R is a weak order defined over $\mathbb{X}$.)

[C] *R is a continuous binary preference relation*

- For any $x \in \mathbb{X}$,
- the sets $U(x) = \{y \in \mathbb{X} : \ yRx\}$ and $L(x) = \{y \in \mathbb{X} : \ xRy\}$
- are closed subsets of $\mathbb{X}$.

[A] *R is an additive binary preference relation*

- For any $x, y, z \in \mathbb{X}$,
- xRy iff $(x + z)R(y + z)$.[14]

[M] *R is a monotonic binary preference relation*

- For any $x, y \in \mathbb{X}$,
- if $x_i R y_i \ \forall i \in \mathbb{S}$,
- then xRy.

---

[14] This axiom makes sense only if the decision maker is risk-neutral. For example, let $n = 2$, and assume $(\alpha, -\alpha)I(-\alpha, \alpha)$. For $x = z = (\alpha, -\alpha)$ and $y = (-\alpha, \alpha)$, additivity would imply $(2\alpha, -2\alpha)I(0, 0)$, which would be sensible only under risk-neutrality.

[N] There exist $x, y \in \mathbb{X}$ such that xPy.

Theorem: R satisfies the axioms W, C, A, M, and N iff there exists a unique probability vector $p \in \Delta^{n-1}$ such that $[xRy]$ iff $[p \cdot x \geq p \cdot y]$.[15]

Elicitation of subjective probabilities

Suppose a decision-maker behaves in accordance with this theorem. We can elicit his subjective probability over any state $i \in \mathbb{S}$ by asking him to report the scalar $\alpha \in [0, 1]$ such that

$$(0, 0, \ldots, 1, \ldots, 0) I (\alpha, \alpha, \ldots, \alpha).$$

Because, this indifference will imply that $p_i = \alpha$.

[15] Δ^{n-1} is the (n-1)-dimensional unit simplex.

Subjective Probabilities II: Savage's Theorem

The Basic Set Up

[] Let $\mathbb{S}$ denote the unstructured set of *states*.
[] Any non-empty subset $A \subseteq S$ is referred to as an *event*.
[] Let $\mathbb{X}$ be the unstructured set of *outcomes*.
[] The function $f : \mathbb{S} \to \mathbb{X}$ is an *act* which specifies the outcome for the decision maker for each state in $\mathbb{S}$. Let $\mathcal{F}$ denote the set of all possible acts.
[] A *constant act* $f_x \in \mathcal{F}_c \subset \mathcal{F}$ leads to the same outcome $x \in \mathbb{X}$ in all states.
[] We are interested in the decision maker's ranking over acts, i.e., the elements of $\mathcal{F}$.
[] R is a binary preference relation over the elements of $\mathcal{F}$. We shall denote the symmetric and asymmetric parts of R by I and P, respectively.
[] For any pair of acts $f, g \in \mathcal{F}$, $[fRg]$ means [*act f is at least as good as act g*].
[] fR_Ag means that given the event $A \subseteq S$, act f is at least as good as act g.
[] $fR_Ag(s)$ means that given $A \subseteq S$, act f is at least as good as the constant act $g(s)$.

The Axioms

[W] R is a transitive and complete binary relation defined over $\mathcal{F}$.

[I_1] The ranking of a pair of acts should be *independent* of the events where they lead to identical outcomes. This is formalized by requiring that what outcomes they lead to in states where they lead to identical outcomes, should not matter.

- For every $f, g, h, h' \in \mathcal{F}$, and any $A \subset S$,
- $[(f, A; h, A^c)R(g, A; h, A^c)]$ iff $[(f, A; h', A^c)R(g, A; h', A^c)]$.

[I_2] The ranking of a pair of acts that differ only on a non-null event and are constant over that event, should be consistent with the ranking of the corresponding constant acts. Implicitly, this amounts to *state-independent* ranking of outcomes.

- For every $f \in \mathcal{F}$, any $f_x, f_y \in \mathcal{F}_c$, and any non-null event $A \subset S$,
- $[(f_x, A; f, A^c)R(f_y, A; f, A^c)]$ iff $[f_xRf_y]$.

[I_3] The inferred (qualitative) relative likelihood of any two events should be *independent* of the pair of constant acts employed in the inference.

- For every $A, B \subset S$, and any $f_x, f_y, f_z, f_w \in \mathcal{F}_c$
- if f_xPf_y and f_zPf_w,
- then $[(f_x, A; f_y, A^c)R(f_x, B; f_y, B^c)]$ iff $[(f_z, A; f_w, A^c)R(f_z, B; f_w, B^c)]$.

[N] There exist $f, g \in \mathcal{F}$ such that fPg.

[$C_{\mathbb{S}}$] Given fPg, we can always find acts f' and g' which are minor perturbations of f and g such that $f'Rg$ and fRg'. It rules out $\mathbb{S}$ to be finite.

- For every $f, g, h \in \mathcal{F}$,
- if fPg,
- then there exists a partition $(S_1, S_2, \ldots, S_n)$ of S such that for all $1 \leq i \leq n$,
- $[(h, S_i; f, S_i^c)Rg]$ and $[fR(h, S_i; f, S_i^c)]$.

[$C_{\mathbb{X}}$]

- For every $f, g \in \mathcal{F}$, any event $A \subset S$,
- if $fR_Ag(s)$ for all $s \in A$, then fR_Ag; and,
- if $g(s)R_Af$ for all $s \in A$, then gR_Af.

Theorem 1: For any $\mathbb{X}$, R satisfies all the seven axioms iff there exists a non-atomic finitely additive probability measure μ over S and a non-constant bounded function $u : \mathbb{X} \to \mathbb{R}$ such that for every $f, g \in \mathcal{F}$,

$$fRg \ \text{ iff } \ \int_{\mathbb{S}} u(f(s))d\mu(s) \ \geq \ \int_{\mathbb{S}} u(g(s))d\mu(s).$$

Theorem 2: For a finite $\mathbb{X}$, R satisfies the axioms W, I_1, I_2, I_3, N, $C_{\mathbb{S}}$ iff there exists a non-atomic finitely additive probability measure μ over S and a non-constant bounded function $u : \mathbb{X} \to \mathbb{R}$ such that for every $f, g \in \mathcal{F}$,

$$fRg \ \text{ iff } \ \int_{\mathbb{S}} u(f(s))d\mu(s) \ \geq \ \int_{\mathbb{S}} u(g(s))d\mu(s).$$

Note 1: The measure μ is unique and the utility function u is unique up to positive linear transformations in both theorems.

Note 2: The theorems essentially lead to (i) state-independent ranking of outcomes and (ii) outcome independent likelihood of events on part of the decision-maker.

Note 3: Axiom I_1 is often referred to as the *sure-thing principle*. Axiom I_2 is referred to as *monotonicity*. Axiom I_3 is the counterpart of I_2 wrt likelihood of events rather than ranking of outcomes. $C_{\mathbb{S}}$ and $C_{\mathbb{X}}$ capture some notion of *continuity*.

Arrow's Impossibility Theorem for $f : \mathbb{L}^N \to \mathbb{L}$

The Basic Set Up

[1] $\mathbb{V}$ is a finite set of voters with $|\mathbb{V}| = N \geq 2$.
[2] $\mathbb{A}$ is a finite set of candidates with $|\mathbb{A}| = M \geq 3$.
[3] $\mathbb{L}$ is the set of all linear orders (strict rankings) on $\mathbb{A}$ with $|\mathbb{L}| = (M!)$.
[4] L denotes an element of the set $\mathbb{L}$. L_i will denote the ranking provided by voter $i \in \mathbb{V}$.
[5] For any $a, b \in \mathbb{A}$, $[aL_ib]$ means [*a is above b in the individual ranking provided by* $i \in \mathbb{V}$].
[6] $L^N \in \mathbb{L}^N$ denotes a possible collection of N rankings. Note: $L^N \equiv (L_1, L_2, \ldots, L_n)$.
[7] $\mathbb{L}^N$ denotes the set of all possible collective rankings with $|\mathbb{L}^N| = (M!)^N$.
[8] $f : \mathbb{L}^N \to \mathbb{L}$ denotes a Social Choice Rule (SCR) that uses the individual rankings provided by the N voters as input and generates a *social ranking* of the candidates as the output. It must produce an output for each of the $(M!)^N$ possible inputs.
[9] $[af(L^N)b]$ means [*a is above b in the social ranking corresponding to the input* L^N].
[10] $f : \mathbb{L}^N \to \mathbb{L}$ is a *Dictatorial* (D) SCR if $\exists! i \in \mathbb{V}$ such that

- for every input $L^N \in \mathbb{L}^N$ and every pair of alternatives $a, b \in \mathbb{A}$,
- $[af(L^N)b]$ *iff* $[aL_ib]$.

[11] The number of possible SCRs is $(M!)^{((M!)^N)}$ (out of which N are dictatorial). The aim is to figure out which of these SCRs satisfy a list of desirable properties formulated as axioms.

[Axiom P] *Pareto SCR*

- For any input $L^N \in \mathbb{L}^N$ and any pair of alternatives $a, b \in \mathbb{A}$,
- if $aL_ib \ \ \forall i \in \mathbb{V}$,
- then $af(L^N)b$.

[Axiom IIA] *Independent of Irrelevant Alternatives SCR*

- For any pair of inputs $L^N, \tilde{L}^N \in \mathbb{L}^N$, and any pair of alternatives $a, b \in \mathbb{A}$.
- if $[aL_ib$ iff $a\tilde{L}_ib$ & bL_ia iff $b\tilde{L}_ia]$ $\forall i \in \mathbb{V}$,
- then $[af(L^N)b$ iff $af(\tilde{L}^N)b$ & $bf(L^N)a$ iff $bf(\tilde{L}^N)a]$.

Theorem: If the SCR $f : \mathbb{L}^N \to \mathbb{L}$ satisfies P and IIA, then it is Dictatorial.

Arrow's Theorem for $f : \mathbb{L}^N \to \mathbb{A}$

The Basic Set Up

[1] $\mathbb{V} = \{1, 2, \ldots, n\}$ is a finite set of voters with $|\mathbb{V}| = N \geq 2$.
[2] $\mathbb{A}$ is a finite set of candidates with $|\mathbb{A}| = M \geq 3$.
[3] $\mathbb{L}$ is the set of all linear orders (strict rankings) on $\mathbb{A}$.
[4] L denotes an element of the set $\mathbb{L}$. L_i will denote the ranking provided by voter $i \in \mathbb{V}$.
[5] For any $a, b \in \mathbb{A}$, $[aL_ib]$ means [*a is above b in the individual ranking provided by $i \in \mathbb{V}$*].
[6] $L^N \in \mathbb{L}^N$ denotes a possible collection of N rankings. Note: $L^N \equiv (L_1, L_2, \ldots, L_n)$.
[7] $\tilde{L}^N$ is a *monotonic transformation of $L^N \in \mathbb{L}^N$ around $a \in \mathbb{A}$* if

- $\forall i \in \mathbb{V}$, and for every $b \in \mathbb{A} - a$,
- if aL_ib,
- then $a\tilde{L}_ib$.

[8] $f : \mathbb{L}^N \to \mathbb{L}$ denotes a Social Choice Rule (SCR) that uses the individual rankings provided by the N voters as input and generates the name of the *winning candidate* as the output.
[9] $f : \mathbb{L}^N \to \mathbb{A}$ is a *Dictatorial* (D) SCF if $\exists! i \in \mathbb{V}$ such that

- for every input $L^N \in \mathbb{L}^N$,
- $f(L^N) \equiv f(L_i, L_{-i}^{N-1}) = a$ *iff* a is at the top of L_i,
- for any $a \in \mathbb{A}$.

The Axioms

[P] *Pareto SCF*

- For any input $L^N \in \mathbb{L}^N$ and any alternative $a \in \mathbb{A}$.
- if L^N is such that a is at the top of L_i $\forall i \in \mathbb{V}$,
- then $f(L^N) = a$.

[MM] *Maskin Monotonic SCF*

- For any pair of inputs $L^N, \tilde{L}^N \in \mathbb{L}^N$ and any alternative $a \in \mathbb{A}$,
- if $\tilde{L}^N$ is a monotonic transformation of L^N around a,
- then [if $f(L^N) = a$, then $f(\tilde{L}^N) = a$].

Theorem 2: For $M \geq 3$, if the SCF $f : \mathbb{L}^N \to \mathbb{A}$ satisfies P and MM, then it is D.

Gibbard-Satterthwaite Theorem

The Basic Set Up

[1] $\mathbb{V} = \{1, 2, \dots, n\}$ is a finite set of voters with $|\mathbb{V}| = N \geq 2$.
[2] $\mathbb{A}$ is a finite set of candidates with $|\mathbb{A}| = M \geq 3$.
[3] $\mathbb{L}$ is the set of all linear orders (strict rankings) on $\mathbb{A}$.
[4] L denotes an element of the set $\mathbb{L}$. L_i will denote the ranking provided by voter $i \in \mathbb{V}$.
[5] For any $a, b \in \mathbb{A}$, [aL_ib] means [*a is above b in the individual ranking provided by* $i \in \mathbb{V}$].
[6] $L^N \in \mathbb{L}^N$ denotes a possible collection of N rankings. Note: $L^N \equiv (L_1, L_2, \dots, L_n)$.
[7] $\tilde{L}^N$ is a *monotonic transformation of* $L^N \in \mathbb{L}^N$ *around* $a \in \mathbb{A}$ if

- $\forall i \in \mathbb{V}$, and for every $b \in \mathbb{A} - a$,
- if aL_ib,
- then $a\tilde{L}_ib$.

[8] $f : \mathbb{L}^N \to \mathbb{L}$ denotes a Social Choice Rule (SCR) that uses the individual rankings provided by the N voters as input and generates the name of the *winning candidate* as the output.
[9] [$f(L^N)L_if(\tilde{L}^N)$] means [*the winning candidate corresponding to input* L^N *is above the winning candidate corresponding to input* $\tilde{L}^N$ *according to* L_i].
[10] $f : \mathbb{L}^N \to \mathbb{A}$ is a *Dictatorial* (D) SCF if $\exists! i \in V$ such that

- for every input $L^N \in \mathbb{L}^N$,
- $f(L^N) \equiv f(L_i, L_{-i}^{N-1}) = a$ *iff* a is at the top of L_i,
- for any $a \in \mathbb{A}$.

The Axioms

[P] *Pareto SCF*

- For any input $L^N \in \mathbb{L}^N$ and any alternative $a \in \mathbb{A}$.
- if L^N is such that a is at the top of L_i $\forall i \in \mathbb{V}$,
- then $f(L^N) = a$.

[MM] *Maskin Monotonic SCF*

- For any pair of inputs $L^N, \tilde{L}^N \in \mathbb{L}^N$ and any alternative $a \in \mathbb{A}$,
- if $\tilde{L}^N$ is a monotonic transformation of L^N around a,
- then [if $f(L^N) = a$, then $f(\tilde{L}^N) = a$].

[CS] *Citizen Sovereign SCF*

- For every alternative $a \in \mathbb{A}$,
- $\exists L^N \in \mathbb{L}^N$ such that,
- $f(L^N) = a$.

[SP] *Strategy Proof SCF*

- For any input $L^N \in \mathbb{L}^N$ and any (other) individual ranking $\tilde{L}_i \in \mathbb{L} - L_i$,
- if the winner at input $(\tilde{L}_i, L_{-i}^{N-1})$ is not the winner at input L^N,
- then the winner at input $(\tilde{L}_i, L_{-i}^{N-1})$ must be below the winner at input L^N according to L_i,
- for every voter in the set of voters.

Result 3: For $M \geq 3$:

[3.1a] If the SCF f is SP, then it will be MM; and

[3.1b] If the SCF f is MM and CS, then it will be PE.

Thus, [3.1a] and [3.1b] imply:

[3.2] If the SCF f is SP and CS, then it will be MM and PE.

We know from Theorem 2 that:

[3.3] If the SCF f is MM and PE, then it will be D.

Hence,

Theorem 3: For $M \geq 3$, if the SCF $f : \mathbb{L}^N \to \mathbb{A}$ satisfies SP and CS, then it is D.

Axiomatization of Pairwise Simple-Majority Rule

The Basic Set Up

[2] $\mathbb{V} \;=\; \{1, 2, \ldots, n\}$ is a finite set of voters with cardinality $N \geq 2$.
[2] $\mathbb{A} \;=\; \{+1, -1\}$ is a finite set of candidates with $|\mathbb{A}| = M = 2$.
[3] $d_i \in \mathbb{D} \;=\; \{+1, 0, -1\}$ denotes a possible input provided by voter $i \in \mathbb{V}$.
[4] $d^N \in \mathbb{D}^N$ denotes a possible collection of N individual inputs (one by each $i \in \mathbb{V}$).
[5] $d^N \in \mathbb{D}^N$ is a N-dimensional vector $(d_1, \ldots, d_i, \ldots, d_n)$ since each d_i is a scalar.
[6] $\mathbb{D}^N$ is the set of all possible inputs and $|\mathbb{D}^N| = 3^N$.
[7] For any input $d^N \in \mathbb{D}^N$, the input where every voter reverses his vote is denoted by $-\mathbf{1} \cdot d^N$.
[8] $\tilde{d}^N$ is called a *permutation* of $d^N \in \mathbb{D}^N$ if both d^N and $\tilde{d}^N$ have same entries but by different voters. Thus, $\tilde{d}^N$ is a permutation of d^N if both have the same *number* of 0's, 1's, and -1's.
[9] $f : \mathbb{D}^N \to \mathbb{D}$ denotes a Social Choice Function (SCF).
[10] $Sign[s()] : \mathbb{D}^N \to \mathbb{D}$ takes a $d^N \in \mathbb{D}^N$ as the input and provides the *sign of the sum of entries* in d^N as the output.
[11] A SCF is the PSMR if $f(d^N) \;=\; Sign[s(d^N)],\ \forall d^N \in \mathbb{D}^N$.

The Axioms

[PR] *Positively Responsive SCF*

- For any pair of inputs $d^N, \tilde{d}^N \in \mathbb{D}^N$,
- if $\exists k \in \mathbb{V}$ such that $[(\tilde{d}_k > d_k)\ \&\ (\tilde{d}_i = d_i\ \forall i \in \mathbb{V} - \{k\})]$,
- then $[f(d^N) \geq 0] \Rightarrow [f(\tilde{d}^N) = +1]$.

[N] *Neutral SCF*

- For any pair of inputs $d^N, \tilde{d}^N \in \mathbb{D}^N$,
- if $[\tilde{d}^N = -\mathbf{1} \cdot d^N]$,
- then $[f(d^N) = -f(\tilde{d}^N)]$.

[A] *Anonymous SCF*

- For any pair of inputs $d^N, \tilde{d}^N \in \mathbb{D}^N$,
- if $[\tilde{d}^N$ is a permutation of $d^N]$
- then $[f(\tilde{d}^N) = f(d^N)]$.

Theorem 4: For $M = 2$, a SCF $f : \mathbb{D}^N \to \mathbb{D}$ satisfies PR, N, and A *iff* it is the PSMR.

One Sided One-to-One Matching Problem

[1] $\mathbb{A}$ is a finite set of agents with cardinality $|\mathbb{A}| = N$.
[2] $\mathbb{H}$ is a finite set of objects with cardinality $|\mathbb{H}| = N$.
[3] $\mathbb{L}$ denotes the finite set of all possible strict rankings over the objects in $\mathbb{H}$.
[4] $\mathbb{L}^N$ denotes the finite set of all possible elements in the N-fold cartesian product of $\mathbb{L}$.
[5] $X^N \equiv (X_1, X_2, \ldots, X_n) \in \mathbb{X}^N$ denotes a possible allocation-vector of the objects to agents such that the element $X_i \in \mathbb{H}$ denotes the object allocated to agent $i \in \mathbb{A}$. Moreover, $|X_i| = 1$ for all $i \in \mathbb{A}$ which implies $\cup_{i\in\mathbb{A}} X_i = \mathbb{H}$, and for any distinct $i, j \in \mathbb{A}$, $X_i \cap X_j = \{\emptyset\}$.
[6] An *allocation mechanism* is a one-to-one and onto function $f : \mathbb{L}^N \to \mathbb{X}^N$ which takes the reported rankings of all the N agents as the input and produces an allocation-vector as the output.
[7] The allocation vector $f(L^N)$ will often be written as $(f_1(L^N), f_2(L^N), \ldots, f_n(L^N))$.
[8] The notation $[f(L^N) = X^N]$ means that [when the mechanism receives the collection of preference rankings $L^N \in \mathbb{L}^N$ as the input, then it produces the allocation vector $X^N \in \mathbb{X}^N$ as the output]. Moreover, $f_i(L^N) = X_i$, $\forall i \in \mathbb{A}$.
[9] The notation $[f_i(L^N) L_i f_i(\tilde{L}^N)]$ means that [agent $i \in \mathbb{A}$ ranks the object he receives when the input is L^N higher than the object he receives when the input is $\tilde{L}^N$ according to preference ranking L_i].

The Axioms

[PE] *Pareto Efficient Mechanism*

- For any input $L^N \in \mathbb{L}^N$, there does not exist any allocation vector $\tilde{X}^N \in \mathbb{X}^N - f(L^N)$, such that,
- $\tilde{X}_i L_i f_i(L^N)$ for at least one $i \in \mathbb{A}$,
- and $f_i(L^N) L_i \tilde{X}_i$ for no $i \in \mathbb{A}$.

[SP] *Strategy Proof Mechanism*

- For any input $(L_i, L_{-i}^{N-1}) \in \mathbb{L}^N$ and any other input $(\tilde{L}_i, L_{-i}^{N-1}) \in \mathbb{L}^N$,
- if $f(\tilde{L}_i, L_{-i}^{N-1}) \neq f(L^N)$,
- then $f_i(L^N) L_i f(\tilde{L}_i, L_{-i}^{N-1})$,
- $\forall i \in \mathbb{A}$.

Theorem: *A mechanism for the standard one-sided matching problem is efficient and strategy proof if and only if it is the Serial Dictatorship mechanism.*

One Sided One-to-Many Matching Problem

[1] $\mathbb{A}$ is a finite set of agents with cardinality $|\mathbb{A}| = N$.
[2] $\mathbb{H}$ is a finite set of objects with cardinality $|\mathbb{H}| = M$.
[3] $P(\mathbb{H})$ denotes the power set (the collection of all subsets) of the set $\mathbb{H}$.
[4] $\mathbb{L}$ is the finite set of all possible strict rankings over the elements in $P(\mathbb{H})$.
[5] $X^N \equiv (X_1, X_2, \ldots, X_n) \in \mathbb{X}^N$ denotes a possible allocation-vector of the objects to agents such that the element $X_i \in P(\mathbb{H})$ denotes the subset of objects in $\mathbb{H}$ allocated to agent $i \in \mathbb{A}$. Moreover, $X_i \in P(\mathbb{H})$, and $0 \leq |X_i| \leq M$ for all $i \in \mathbb{A}$, $\cup_{i \in \mathbb{A}} X_i \subseteq \mathbb{H}$, and for any distinct $i, j \in \mathbb{A}$, $X_i \cap X_j = \{\emptyset\}$.
[6] An *allocation mechanism* is a one-to-one function $f : \mathbb{L}^N \to \mathbb{X}^N$ which takes the reported rankings of all the N agents as the input and produces an allocation-vector as the output.
[7] The allocation vector $f(L^N)$ will often be written as $(f_1(L^N), f_2(L^N), \ldots, f_n(L^N))$.
[8] The notation $[f(L^N) = X^N]$ means that [when the mechanism receives the collection of preference rankings $L^N \in \mathbb{L}^N$ as the input, then it produces the allocation vector $X^N \in \mathbb{X}^N$ as the output]. Moreover, $f_i(L^N) = X_i$, $\forall i \in \mathbb{A}$.
[9] The notation $[f_i(L^N) L_i f_i(\tilde{L}^N)]$ means that [agent $i \in \mathbb{A}$ ranks the set of objects he receives when the input is L^N higher than the set of objects he receives when the input is $\tilde{L}^N$ according to preference ranking L_i].

The Axioms

[PE] *Pareto Efficient Mechanism*

- For any input $L^N \in \mathbb{L}^N$, there does not exist any allocation vector $\tilde{X}^N \in \mathbb{X}^N - f(L^N)$, such that,
- $\tilde{X}_i L_i f_i(L^N)$ for at least one $i \in \mathbb{A}$, and $f_i(L^N) L_i \tilde{X}_i$ for no $i \in \mathbb{A}$.

[SP] *Strategy Proof Mechanism*

- For any input $(L_i, L_{-i}^{N-1}) \in \mathbb{L}^N$ and any other input $(\tilde{L}_i, L_{-i}^{N-1}) \in \mathbb{L}^N$,
- if $f(\tilde{L}_i, L_{-i}^{N-1}) \neq f(L^N)$, then $f_i(L^N) L_i f(\tilde{L}_i, L_{-i}^{N-1})$, $\forall i \in \mathbb{A}$.

[NB] *Non-bossy Mechanism*

- For any input $(L_i, L_{-i}^{N-1}) \in \mathbb{L}^N$ and any other input $(\tilde{L}_i, L_{-i}^{N-1}) \in \mathbb{L}^N$,
- if $f_i(L^N) = f_i(\tilde{L}_i, L_{-i}^{N-1})$, then $f(L^N) = f(\tilde{L}_i, L_{-i}^{N-1})$, $\forall i \in \mathbb{A}$.

Theorem: *An allocation mechanism for the one-sided matching problem with multiple assignments is efficient, strategy proof, and non-bossy if and only if it is the Sequential Dictatorship mechanism.*

Useful References

[] Arrow, K. J. *Social Choice and Individual Values.* Wiley, 1951.

[] Bourbaki, Nicholas. "The Architecture of Mathematics." *American Mathematical Monthly*, 19(1), 1950, pp. 175-198.

[] Debreu, G. *Theory of Value.* Yale Univ. Press, 1959.

[] Dietrich, F. "Judgment Aggregation: (Im)possibility Theorems." *Journal of Economic Theory*, 126(1), 2006, pp. 286-298.

[] Gaertner, W. *A Primer in Social Choice Theory.* Oxford Univ. Press, 2006.

[] Gibbard, A. "Manipulation of Voting Schemes." *Econometrica*, 41, 1973, pp. 587-602.

[] Gilboa, I. *Theory of Decision Under Uncertainty.* Cambridge Univ. Press, 2009.

[] Gilboa, I. and Schmeidler, D. "Inductive Inference: An Axiomatic Approach." *Econometrica*, 71, 2003, pp. 1-26.

[] Gowers, T. "Language and Grammar of Mathematics." *The Princeton Companion to Mathematics*, Princeton Univ. Press, 2008.

[] Halmos, P. R. *Naive Set Theory.* Springer-Verlag, 1974.

[] List, C. and Pettit, P. "Aggregating Sets of Judgments: An Impossibility Result." *Economics and Philosophy*, 18, 2002, pp. 89110.

[] Manzini, P. and Mariotti, M. "Sequentially Rationalizable Choice." *American Economic Review*, 97(5), 2007, pp. 1824-1839.

[] Mas-Colell, A., Whinston, M. D., and Green, J. R. *Microeconomic Theory.* Oxford Univ. Press, 1995.

[] Ok, Efe O. *Real Analysis With Applications.* Princeton Univ. Press, 2006.

[] Nagel, E. and Newman, J. *Godel's Proof.* New York Univ. Press, 2008.

[] Palacios-Huerta, I. and Volij, O. "The Measurement of Intellectual Influence." *Econometrica*, 72(3), 2004, pp. 963-977.

[] Rubinstein, A. and Salant, Y. "Two Comments on the Principle of Revealed Preference." Working Paper, 2008.

[] Satterthwaite, M. "Strategy-Proofness and Arrow's Conditions: Existence and Correspondence Theorems for Voting procedures and Social Welfare Functions." *Journal of Economic Theory*, 10, 1975, pp. 187-217.

[] Savage, L. J. *The Foundations of Statistics.* Dover Publications, 1954.

[] Sen, A. K. "On Weights and Measures: Informational Constraints in Social Welfare Analysis." *Econometrica*, 45, pp. 1539-1572.

[] Stokey, N. L. and Lucas, R. L. *Recursive Methods in Economic Dynamics.* Harvard Univ. Press, 1989.

[] Thomson, W. "On the Axiomatic Method and its Recent Applications to Game Theory and Resource Allocation." *Social Choice and Welfare*, 18, 2001, pp. 327-386.

[] Velupillai, V. *Computational Foundations for Economics.* Rouletdge, 2010.

[] von-Neumann, J. and Morgenstern, O. *Theory of Games and Economic Behavior.* Princeton Univ. Press, 1944.

[] Weintraub, E. R. *How Economics Became a Mathematical Science.* Duke Univ. Press, 2002.

[] Woeginger, G. J. "An Axiomatic Characterization of the Hirsch-index." *Mathematical Social Sciences*, 56(2), 2008, pp. 157-183.